金融学前沿系列

QUANTITATIVE FINANCIAL ECONOMICS

Stocks, Bonds & Foreign Exchange (Second Edition)

数量金融经济学

（第二版）

[英] Keith Cuthbertson
Dirk Nitzsche 著

朱 波 译

WILEY

西南财经大学出版社

金融学前沿系列

F830/201
2008

QUANTITATIVE FINANCIAL ECONOMICS

Stocks, Bonds & Foreign Exchange (Second Edition)

数量金融经济学

（第二版）

[英] **Keith Cuthbertson**
Dirk Nitzsche 著

朱 波 译

WILEY

西南财经大学出版社

Quantitative Financial Economics: Stocks, Bonds and Foreign Exchange. Second Edition / Keith Cuthbertson and Dirk Nitzsche.
All Rights Reserved. Authorised translation from the English language edition published by John Wiley & Sons, Ltd.

图书在版编目(CIP)数据

数量金融经济学/(英)Cuthbertson，K.；（英）Nitzsche，D. 著；朱波译．
—2版．—成都：西南财经大学出版社，2008.4
ISBN 978-7-81088-876-9

Ⅰ．数… Ⅱ．①C…②N…③朱… Ⅲ．金融学 Ⅳ．F830

中国版本图书馆 CIP 数据核字(2007)第 172008 号

数量金融经济学(第二版)
[英]Keith Cuthberston Dirk Nitzsche 著
朱波 译

责任编辑：李霞湘
封面设计：杨红鹰
责任印制：封俊川

出版发行：	西南财经大学出版社(四川省成都市光华村街 55 号)
网　　址：	http://www.xcpress.net
电子邮件：	xcpress@mail.sc.cninfo.net
邮政编码：	610074
电　　话：	028-87353785 87352368
印　　刷：	四川森林印务有限责任公司
成品尺寸：	185mm×260mm
印　　张：	32
字　　数：	680 千字
版　　次：	2008 年 4 月第 2 版
印　　次：	2008 年 4 月第 1 次印刷
印　　数：	1—2000 册
书　　号：	ISBN 978-7-81088-876-9
定　　价：	69.80 元

1. 如有印刷、装订等差错，可向本社营销部调换。
2. 版权所有，翻印必究。
3. 本书封底无本社数码防伪标志，不得销售。

译者序

Keith Cuthbertson 和 Dirk Nitzsche 撰写的《数量金融经济学》（第二版）对金融经济学的新近发展和实证方法进行了系统全面的介绍。全书不仅对金融经济学概念和思想背后的直觉进行了阐述，还对用数学和统计学方法进行实证的思想进行了叙述，部分领域涉及金融经济学研究的前沿。阅读此书可以使读者很快达到学术前沿，也能为读者在金融衍生工具定价、金融工程和金融风险管理等领域的学习和研究奠定坚实的基础。本书是金融经济学、实证金融学和金融数量分析等课程的标准教科书。

本书内容覆盖了资产定价、投资决策和特殊策略分析的主要理论思想。主要内容包括金融中的基本概念、金融中的统计方法、有效市场假说及检验、资产收益的可预测性、线性因子模型及检验和应用、CCAPM 和 SDF、跨期资产配置理论与实证、期限结构理论与实证、金融异象与行为金融、外汇市场和市场微观结构理论等，内容丰富，难度适中。

本书可以用作金融学专业硕士研究生的数量分析教材，也是金融工程专业和数理金融专业硕士研究生的重要参考书，对金融数量分析感兴趣的高年级本科生和金融从业人员而言，本书也是非常有用的读物。本书也可以用作金融经济学课程的辅助读物。

在翻译过程中，西南财经大学金融学院的研究生李鸿良、陈聪慧、金晨、郭小平和雷海林为本书的文字输入做了大量的工作，在此表示感谢。感谢西南财经大学出版社的方英仁总编，正是在他的组织和协调下，本书才得以顺利出版。感谢本书的责任编辑李霞湘女士所给予的大力支持和帮助。

由于译者学术水平有限，加上时间仓促，本书中错误之处在所难免，敬请读者批评指正。

译者

2008 年 3 月于成都

序言

大量的来访邮件和本书第一版的售罄都表明这是一个NPV为正的投资项目，这种状况令人鼓舞。但是，如果不能持久，那么对作者来说仅有现金回报是远远不够的。近年来，我的妻子认为我每小时努力工作的报酬赶不上一个熟练的水管工人（她是对的）。但是，她也得考虑正的外部性和实物期权理论——不管怎样，这是我的反驳观点。尽管我属于风险厌恶型，尽管我的长期净债务具有时变性，我还是无意放弃我的日常工作。

有人邀请你参加晚宴，为你奉上美酒，然后要求你出版书稿的第二版，如果你在这个关键时刻毫不含糊地说“好”，那么你就上当了。原因很明显。这个领域在过去十年中出版的著作可以充满整个“阿尔伯特大厅”，当然电子版的不算。研究结果和第一稿是最棒的，接下来的稿子就要差一些，到最后就只剩下痛苦的煎熬了——简直就是“土拨鼠的日子”。对我而言，写一本书是十分困难的。然而，幸运的是，在本书第二版的写作过程中，我能很好地鼓励我的优秀合作者Dirk Nitzsche。

对那些购买了本书第一版的读者而言，扫描一下扉页上的内容简介就能发现，第二版65%的内容是新增的，所保留的第一版的内容在这一版中也进行了修正。我们希望我们所选择的主题具有连贯性，有趣而又范围宽广。对于非法影印本书第一版的那些朋友而言，我很钦佩你们为海外援助所做的贡献，尽管你们的行为不是自觉行为。但是，对这个尽心修改的版本，我们“希望你们能够付费”。

为谁而作？

本书可以作为金融专业硕士和金融经济学专业硕士的数量分析教材，对博士生而言本书也是非常有用的读物。你所需要的基础知识是本科层次的一些基础金融理论（伴随一些数学基础知识）和现代时间序列计量经济学的相关知识。对金融从业人士而言，如果想知道学术研究是否具有实践价值（答案是“有时有价值”），那么也应该深入阅读一些细节。至少，你能分辨出一个糟糕领袖从象牙塔走向市场能否从容应对挑战。

本书主要内容覆盖了资产定价、投资策略确定（使用离散时间分析）和特殊策略分析的主要理论思想。我们给出了具有说服力的实证结果，尽管这会有所遗漏（我们有时期望有所遗漏）。重点是金融和经济学概念背后的直觉，我们也用数学和统计学方法来对这些思想进行严谨分析。我们的感觉是，这些材料可以使读者进入这些领域的研究前沿。我们也希望本书能鼓励读者在金融衍生工具定价、金融工程和风险管理领域进行探索。

我们希望本书能为你带来愉悦，能使你的财富陡增——谁知道呢，也可能这是你拥有美妙友谊的开始。经过我们过去的努力，事实上我们已经尽了最大努力，到目前为止，我们的目的是为了理解一些东西，而不仅仅是为了出版。本书能否增加社会福利，这个问题只有时间和你才能给出答案。

Keith Cuthbertson
Dirk Nitzsche
2004年10月

致谢

我们特别感谢CASS商学院的同事们，他们为本书多个主题的实践见解提供了帮助。我们也要感谢帝国理工大学Tanaka商学院的同事们。这两所大学和法国波尔多第四大学、墨西哥国立自治大学、柏林自由大学的硕士和博士生是本书非常优秀的、富有批判性的、非常细心的“试验者”。

我们要特别感谢那些阅读过本书初稿的人士：Mike Artis，Don Bredin，Francis Breedon，Ian Buckley，Alec Chrystal，David Barr，Lara Cathcart，Ales Cerny，Lina El－Jahel，Luis Galindo，Stephen Hall，Richard Harris，Simon Hayes，Stuart Hyde，Turalay Kenc，David Miles，Michael Moore，Kerry Patterson，Mark Salmon，Lucio Sarno，Peter Smith，Mark Taylor，Dylan Thomas和Mike Wichens。

奥胡斯大学的Tom Engsted和埃克塞特大学的Richard Harris指出了本书第一版中一些错误，我们希望第二版能对这些错误进行更正。此外，Niall O’Sullivan允许我们在本书第9章共同基金绩效部分中使用他的部分PhD资料，Ales Cerny很早就同意我们使用《金融中的数学技巧》（普林斯顿大学出版社2004年版）一书的书稿。

我们要感谢帝国理工大学的Yvonne Doyle，这位正努力成为希腊专家的女士非常专业地录入了本书的大量书稿。

目录

第 1 章　金融中的基本概念 …… (1)
1.1　股票、债券和实物资产的收益 …… (1)
1.2　折现现值（DPV） …… (5)
1.3　效用与无差异曲线 …… (9)
1.4　资产需求 …… (14)
1.5　无差异曲线与跨期效用函数 …… (18)
1.6　投资决策与最优消费 …… (20)
1.7　小结 …… (23)
附录：均值方差模型与效用函数 …… (24)

第 2 章　金融中的统计基础 …… (25)
2.1　对数正态分布与 Jensen 不等式 …… (25)
2.2　单位根、随机游走与协整 …… (26)
2.3　蒙特卡罗模拟与自助法 …… (28)
2.4　贝叶斯学习 …… (33)
2.5　小结 …… (36)

第 3 章　有效市场假说 …… (37)
3.1　概述 …… (37)
3.2　EMH 的含义 …… (39)
3.3　数学期望、鞅与公平赌博 …… (41)
3.4　检验 EMH …… (44)
3.5　使用调查数据 …… (45)
3.6　小结 …… (48)
附录：交叉方程限制 …… (48)

第 4 章　股票收益可预测吗 …… (50)
4.1　一个世纪的收益 …… (50)
4.2　简单模型 …… (57)
4.3　单变量检验 …… (59)
4.4　多变量检验 …… (66)
4.5　协整与误差修正模型 …… (69)
4.6　非线性模型 …… (71)
4.7　马尔可夫转换模型 …… (74)
4.8　交易策略可以获利吗? …… (76)
4.9　小结 …… (79)

第 5 章　均值方差组合理论与资本资产定价模型 …… (80)
5.1　概述 …… (80)
5.2　均值方差模型 …… (83)
5.3　资本资产定价模型 …… (92)
5.4　β 与系统风险 …… (93)
5.5　小结 …… (96)

第 6 章　分散化国际投资组合 …… (98)
6.1　均值方差模型的数学推导 …… (99)
6.2　国际分散化 …… (105)
6.3　实践中的均值方差优化问题 …… (108)
6.4　小结 …… (112)
附录 I：有效前沿和资本市场线 …… (112)
附录 II：市场投资组合 …… (115)

第 7 章　绩效度量、CAPM 与 APT …… (117)
7.1　绩效度量 …… (117)
7.2　CAPM 模型的扩展 …… (122)
7.3　单因子模型 …… (124)
7.4　套利定价理论 …… (125)
7.5　小结 …… (129)

第 8 章　CAPM 和 APT 的经验证据 …… (130)
8.1　CAPM：时间序列检验 …… (130)
8.2　CAPM：横截面检验 …… (131)
8.3　CAPM、多因子模型和 APT …… (134)
8.4　小结 …… (140)
附录：Fama - MacBeth 两阶段回归 …… (140)

第 9 章　线性因子模型的应用 …… (142)
9.1　事件研究 …… (142)
9.2　共同基金的绩效 …… (145)
9.3　共同基金之“星” …… (157)
9.4　小结 …… (169)

第 10 章　价值评估模型与资产收益 …… (170)
10.1　理性定价公式 …… (170)
10.2　理性定价公式的特殊情形 …… (172)
10.3　时变预期收益 …… (173)
10.4　小结 …… (175)

第 11 章　股票价格的波动率 …… (177)
11.1　Shiller 波动率检验 …… (178)

11.2 波动率检验与平稳性 …………………………………… (181)
11.3 比索问题与方差边界检验 …………………………………… (185)
11.4 波动率检验与回归检验 …………………………………… (186)
11.5 小结 …………………………………… (187)
附录：LeRoy - Porter 检验和 West 检验 …………………………………… (187)

第 12 章 股票价格：向量自回归方法 …………………………………… (190)
12.1 收益与理性定价公式的线性化 …………………………………… (190)
12.2 实证结果 …………………………………… (195)
12.3 持续性与波动率 …………………………………… (202)
12.4 小结 …………………………………… (205)
附录：收益、方差分解与持续性 …………………………………… (205)

第 13 章 SDF 模型与 CCAPM 模型 …………………………………… (210)
13.1 以消费为基础的资本资产定价模型 …………………………………… (210)
13.2 CCAPM 模型与“标准” CAPM 模型 …………………………………… (214)
13.3 价格与协方差 …………………………………… (217)
13.4 理性定价公式与 SDF 模型 …………………………………… (218)
13.5 因子模型 …………………………………… (219)
13.6 小结 …………………………………… (220)
附录：联合对数正态分布与幂效用函数 …………………………………… (220)

第 14 章 CCAPM 模型：证据与扩展 …………………………………… (223)
14.1 在 CCAPM 模型中资产收益是否可预测？ …………………………………… (223)
14.2 股权溢价之谜 …………………………………… (226)
14.3 CCAPM 模型欧拉方程的检验 …………………………………… (229)
14.4 SDF 模型的扩展 …………………………………… (232)
14.5 习惯形成 …………………………………… (239)
14.6 股权溢价：更进一步的解释 …………………………………… (242)
14.7 小结 …………………………………… (244)
附录：Hansen - Jagannathan 边界 …………………………………… (244)

第 15 章 跨期资产配置：理论 …………………………………… (246)
15.1 两期模型 …………………………………… (246)
15.2 多期模型 …………………………………… (250)
15.3 期望收益的 SDF 模型 …………………………………… (254)
15.4 小结 …………………………………… (255)
附录 I：消费 - 投资组合问题的包络条件 …………………………………… (255)
附录 II：对数效用函数情形的解 …………………………………… (256)

第 16 章 跨期资产配置：实证 …………………………………… (259)
16.1 退休与随机收入 …………………………………… (259)
16.2 多种风险资产 …………………………………… (262)

16.3　不同偏好 …… (263)
16.4　期限效应与不确定性 …… (265)
16.5　择时交易与不确定性 …… (267)
16.6　随机参数 …… (268)
16.7　稳健性 …… (269)
16.8　小结 …… (270)
附录：参数不确定性与贝叶斯定理 …… (270)

第17章　理性泡沫与学习 …… (273)
17.1　理性泡沫 …… (273)
17.2　理性泡沫的检验 …… (276)
17.3　内在泡沫 …… (278)
17.4　学习 …… (281)
17.5　小结 …… (288)

第18章　行为金融与异像 …… (290)
18.1　主要思想 …… (290)
18.2　信念与偏好 …… (293)
18.3　噪声交易者的存活 …… (295)
18.4　异像 …… (297)
18.5　公司金融 …… (307)
18.6　小结 …… (308)

第19章　行为模型 …… (309)
19.1　简单模型 …… (309)
19.2　噪声交易者行为的优化模型 …… (311)
19.3　Shleifer - Vishny 模型：短期主义 …… (315)
19.4　传染 …… (317)
19.5　信念与预期 …… (319)
19.6　动量与信息观察者 …… (320)
19.7　风格投资 …… (322)
19.8　前景理论 …… (325)
19.9　小结 …… (332)
附录 I DeLong 等的噪声交易者模型 …… (333)
附录 II：Shleifer - Vishny 的短期主义模型 …… (334)

第20章　期限结构理论 …… (336)
20.1　价格、收益与 RVF …… (337)
20.2　期限结构理论 …… (339)
20.3　预期假说 …… (342)
20.4　小结 …… (343)

第 21 章　预期理论：从理论到实证 ……………………………… (345)
21.1　EH 的各种表示 ……………………………… (345)
21.2　VAR 方法 ……………………………… (348)
21.3　时变期限溢价——VAR 方法 ……………………………… (352)
21.4　小结 ……………………………… (353)

第 22 章　期限结构的经验证据 ……………………………… (355)
22.1　数据与协整分析 ……………………………… (355)
22.2　方差边界检验 ……………………………… (357)
22.3　单方程检验 ……………………………… (358)
22.4　预期假说：案例研究 ……………………………… (361)
22.5　以前的研究 ……………………………… (368)
22.6　小结 ……………………………… (370)

第 23 章　随机折现因子与彷射期限结构模型 ……………………………… (372)
23.1　SDF 模型 ……………………………… (372)
23.2　单因子彷射模型 ……………………………… (375)
23.3　多因子彷射模型 ……………………………… (376)
23.4　小结 ……………………………… (377)
附录 I：期限结构 SDF 模型中的数学 ……………………………… (378)
附录 II：单因子彷射模型 ……………………………… (379)

第 24 章　外汇市场 ……………………………… (381)
24.1　汇率制度 ……………………………… (381)
24.2　购买力平价与一价定律 ……………………………… (383)
24.3　抵补利率平价 ……………………………… (388)
24.4　非抵补利率平价 ……………………………… (388)
24.5　远期汇率无偏性 ……………………………… (389)
24.6　实际利率平价 ……………………………… (389)
24.7　小结 ……………………………… (390)
附录：PPP 与工资 - 价格螺旋 ……………………………… (390)

第 25 章　检验 CIP、UIP 和 FRU ……………………………… (392)
25.1　抵补套利 ……………………………… (392)
25.2　非抵补利率平价 ……………………………… (395)
25.3　远期汇率无偏性 ……………………………… (397)
25.4　检验 FRU：VAR 方法 ……………………………… (401)
25.5　比索问题与学习 ……………………………… (404)
25.6　小结 ……………………………… (406)

第 26 章　外汇风险溢价建模 ……………………………… (407)
26.1　FRU 回归中 $\beta<1$ 的含义 ……………………………… (407)
26.2　CCAPM 模型 ……………………………… (408)

26.3 外汇收益的仿射模型 …………………………………………… (411)
26.4 FRU 与货币先行模型 …………………………………………… (412)
26.5 小结 …………………………………………………………… (416)

第 27 章 汇率与基本经济因素 …………………………………… (417)
27.1 货币模型 ……………………………………………………… (417)
27.2 模型检验 ……………………………………………………… (423)
27.3 新开放经济宏观经济学 ………………………………………… (428)
27.4 小结 …………………………………………………………… (429)

第 28 章 市场风险 ……………………………………………… (430)
28.1 Var 的度量 …………………………………………………… (430)
28.2 资产映射：计算的简化 ………………………………………… (435)
28.3 非参数方法 …………………………………………………… (438)
28.4 蒙特卡罗模拟 ………………………………………………… (439)
28.5 其他方法 ……………………………………………………… (442)
28.6 小结 …………………………………………………………… (444)
附录 I：蒙特卡罗分析与 Var …………………………………… (444)
附录 II：单一指数模型 ………………………………………… (446)

第 29 章 波动率与市场微观结构 ………………………………… (448)
29.1 波动率 ………………………………………………………… (448)
29.2 波动率的影响因素 …………………………………………… (450)
29.3 多变量 GARCH 模型 ………………………………………… (456)
29.4 市场微观结构——外汇交易 …………………………………… (460)
29.5 调查数据与预期 ……………………………………………… (461)
29.6 技术交易法则 ………………………………………………… (465)
29.7 小结 …………………………………………………………… (466)

参考文献 ……………………………………………………… (468)

推荐读物 ……………………………………………………… (496)

第 1 章

金融中的基本概念

目的

* 考虑纯贴现债券、付息债券和股票收益的不同度量方法。
* 使用贴现现值（DPV）来对资产进行定价。
* 说明如何使用效用函数来刻画风险规避型偏好，如何根据单期效用最大化来导出资产的需求函数。
* 在两期问题中，导出实际投资和消费的最优水平。

本章的目的是快速地梳理金融文献中要用到的一些基本分析工具。我们没有完全覆盖所有的主题，只是在相当直观的层面来展开讨论。

1.1　股票、债券和实物资产的收益

在金融领域，许多理论研究考虑的是复利收益率或复利利率，尽管利率的市场报价使用的是“单利利率”。例如，就每 6 个月支付一次的 5% 的利率而言，市场报价的单利利率为每年 10%。然而，如果投资者对两个半年账户进行滚动投资且利率为常数，那么他实际获得的“复利”年利率或“真实”年利率或“有效”年利率为$1.05^2 = 1.1025$ 或 10.25%。有效年收益率大于单利利率，因为在有效年收益情形中，投资者获得了“利息的利息（利滚利）”。

现在考察如下问题：在复利利率的计息频率改变后，我们应如何计算一笔投资的最终价值。显然，如果每月计算一次利息，那么每年 10% 的市场报价利率在年末的总利息（复利利息）比每年计算一次利息的年末总利息（单利利息）要多。

考虑年利率为 R、投资期限为 n 年、金额为 A 美元的一笔投资（这里的 R 用小数来进行表示）。如果只是在每年的年末计息，那么这笔投资在 n 年后的未来价值为 FV_n，

$$\mathrm{FV}_n = A\ (1 + R)^n \text{美元} \tag{1}$$

然而，如果每年支付 m 次利息，那么 n 年后这笔投资的最终价值为

$$\mathrm{FV}_n^m = A\ (1 + R/m)^{mn} \text{美元} \tag{2}$$

经常称 R/m 为每期利率（period interest）。随着计息频率的增加，利率就演变为“连续复利”（continuous compounded）利率。可以证明，这笔投资的连续复利最终价值为

$$\mathrm{FV}_n^c = Ae^{R_c n} \text{美元} \tag{3}$$

这里的 R_c 是每年连续复利利率。例如，如果市场报价（单利）利率为每年 10%，那么表 1 给出的是不同计息频率 m 下 100 美元现值在一年末（$n = 1$）的最终价值。如果每天计息一次且 R 为每年 10%，那么根据（2）式，一年后的最终价值为 110.5155 美元。如果假设 $R_c =$

10%，那么 $FV_n^c = 100e^{0.10\times1} = 100.5171$ 美元。因此，每天计息一次所得到的最终价值几乎等于使用连续复利利率计算的最终价值（表 1 的最后两行）。

表 1　　　　计息频率

计息频率	100 美元现值在年末的最终价值（R 为每年 10%）
每年（$m=1$）	110.00
每季度（$m=4$）	110.38
每周（$m=52$）	110.51
每天（$m=365$）	110.5155
连续（$n=1$）	110.5171

现在考虑如何在单利利率、每期利率、有效年利率和连续复利利率之间进行转换。假设一笔投资的利率是每季度 2% 的每期利率，市场报价利率通常是每年 8% 的单利利率。$A=100$ 美元的现值在年末会增值到

$$A\ (1+R/m)^m = 100\ (1+0.08/4)^4 = 108.24\text{美元} \tag{4}$$

有效年利率 R_e 为 8.24%，因为 $100\ (1+R_e) = 108.24$ 美元。因为利滚利，所以 R_e 大于单利利率。每年支付 m 次的单利利率 R 与有效年利率 R_e 之间的关系为

$$(1+R_e) = (1+R/m)^m \tag{5}$$

利用（5）式，可以将每期利率转化为有效利率，也可以将有效利率转化为每期利率。例如，如果有效利率为每年 12%，那么每季度进行一次支付的利率可由 $1.12 = (1+R/4)^4$ 来给出，于是

$$R = [(1.12)^{1/4} - 1] \times 4 = 0.0287 \times 4 = 11.48\% \tag{6}$$

因此，如果每季度计息一次，那么每年 11.48% 的单利利率等价于每年 12% 的有效利率。

使用类似程序，我们能在单利利率 R 与等价的连续复利利率 R_c 之间进行转换，单利利率的计息频率为每年 m 次。进行这些计算的一个原因是，债券（期货和期权）定价的许多复杂理论都要使用连续复利利率。

假设一年支付 m 次的利率 R 是已知的，我们希望计算出 R_c 的值。因为 A 美元投资在 n 年后的最终价值（无论用那种利率进行计算）都必须相等，所以我们有

$$Ae^{R_c n} = A\ (1+R/m)^{mn} \tag{7}$$

因此·

$$R_c = m\ln[1+R/m] \tag{8}$$

此外，如果连续复利利率 R_c 是已知的，那么使用上述方程就能计算出每年计息 m 次的年单利利率 R，

$$R = m(e^{R_c/m} - 1) \tag{9}$$

我们用一个具有说服力的例子来对各种利率之间的上述关系进行总结。假设每 6 个月计息一次的每期利率为 5%（$m=2$，$R/2=0.05$）。市场报价的“单利利率”为每年 10%。100 美元投资在一年后的价值变为 $100\ [1+(0.10/2)]^2 = 110.25$ 美元（使用方程(2)）。显然，有效年利率为每年 10.25%。假设我们希望将 $R=0.10$ 的单利利率转换为等价的连续复利利率。使用（8）式并取 $m=2$ 就可以看出 $R_c = 2\ln\ (1+0.10/2) = 0.09758$（即每年 9.758%）。当然，如果利率是每年 9.758% 的连续复利利率，那么今天 100 美元的投资在一年之后（$n=1$）会增值到 $100e^{R_c n} = 110.25$ 美元。

算术平均与几何平均

假设连续三期的价格分别为 $P_0=1$，$P_1=0.7$ 和 $P_2=1$，相应的（每期）收益率分别为

$R_1=-0.30$（-30%），$R_2=0.42857$（42.857%）。因此，算术平均收益为 $\bar{R}=(R_1+R_2)/2=6.4285\%$。然而，进行如下假设很有可能是错误的：如果你拥有的初始财富为 $W_0=100$ 美元，那么2期后最终财富就变成 $W_2=(1+\bar{R})^2W_0=113.27026$ 美元。观察价格序列就可以明显看出，你的财富在 $t=0$ 和 $t=2$ 之间没有任何改变：

$$W_2=W_0[(1+R_1)(1+R_2)]=100(0.70)(1.42857)=100\text{ 美元}$$

将几何平均收益定义为

$$(1+\bar{R}_g)^2=(1+R_1)(1+R_2)=1$$

从而 $\bar{R}_g=0$，它正确反映了“财富组合”在 $t=0$ 到 $t=2$ 之间的收益，

$$R_w(0\to2)=(W_2/W_0)-1=0$$

一般而言，几何平均收益可以定义为

$$(1+\bar{R}_g)^n=(1+R_1)(1+R_2)\cdots(1+R_n) \tag{10}$$

我们总可以将财富写成

$$W_n=W_0(1+\bar{R}_g)^n$$

除非（每期）收益 R_t 为常数，否则几何平均收益总是小于算术平均收益。例如，当使用年收益 R_t 时，美国在1802—1997年期间股票价值加权指数的几何平均收益为每年7%，远远低于每年8.5%的算术平均收益（Siegel（1998））。

如果收益序列不相关，$R_t=\mu+\varepsilon_t$，$\varepsilon_t\sim iid(0,\sigma^2)$，那么算术平均收益是未来任意年份收益的最佳预测值。在很长的持有期中，最佳预测值也可以用算术平均复利收益来进行计算，即 $(1+\bar{R})^n$。不幸的是，在较长时期中，实践领域并不应用后面一个直观而又简单的结论，因为长期股票收益不是独立同分布序列。

在我们的简单例子中，如果两个收益率重复出现，那么收益序列负相关（即 -30% 和 $+42.8\%$ 在以后各期中交替出现）。在这种情形中，对长期收益进行预测要求使用几何平均复利收益，即 $(1+\bar{R}_g)^n$。有证据表明，长期股票收益具有“温和”的均值反转特征（即收益具有一定程度的序列负相关特征），从而算术平均收益会高估预期未来收益。因此，我们最好将几何平均收益作为未来平均收益的预测值。

长期

（每期）收益为 $(1+R_1)=P_1/P_0$。在跨期模型中，我们经常需要将最终财富表示为

$$W_n=W_0(1+R_1)(1+R_2)\cdots(1+R_n)$$

一种替代表示是

$$\begin{aligned}\ln(W_n/W_0)&=\ln(1+R_1)+\ln(1+R_2)+\cdots+\ln(1+R_n)\\&=R_{c1}+R_{c2}+\cdots+R_{cn}\\&=\ln(P_n/P_0)\end{aligned}$$

这里的 $R_{ct}\equiv\ln(1+R_t)$ 是连续复利利率。注意，括号中的项等于 $\ln(P_n/P_0)$。于是，

$$W_n=W_0\exp(R_{c1}+R_{c2}+\cdots+R_{cn})=W_0(P_n/P_0)$$

因为连续复利利率具有可加性，所以我们能在 $t=0$ 到 $t=n$ 的整个时期上将（总的连续复利）收益定义为

$$R_c(0\to n)\equiv(R_{c1}+R_{c2}+\cdots+R_{cn})$$

$$W_n=W_0\exp[R_c(0\to n)]$$

现在，我们将连续复利收益与几何平均收益“联系”起来。由（10）可以得到

$$\ln(1+\bar{R}_g)^n=R_{c1}+R_{c2}+\cdots+R_{cn}\equiv R_c(0\to n)$$

因此

$$W_n=W_0\exp[\ln(1+\bar{R}_g)^n]=W_0(1+\bar{R}_g)^n$$

这与前面所得到的结论一致。

名义收益和实际收益

许多资产定价模型关注的是实际收益而不是名义收益。实际收益是相对于商品和服务购买力而言的投资（百分比）收益率。实际收益（比方说每年3%）意味着你的初始投资允许你在年末多购买3%固定篮子（basket）的国内物品（例如，对英国居民而言的哈罗德Hamper）。

如果你在$t=0$时拥有名义财富W_0，那么你的实际财富就为$W_0^r=W_0/P_0^g$，这里的P^g是商品和服务价格指数。如果R是你财富的名义（比例）收益，那么你在1年末的名义财富为$W_0(1+R)$，相应的实际财富为

$$W_1^r \equiv \frac{W_1}{P_1^g}=\frac{(W_0^r P_0^g)(1+R)}{P_1^g}$$

因此，你实际财富的增加速度（等价地说，你的（比例）实际收益）为

$$(1+R^r)\equiv W_1^r/W_0^r=(1+R)/(1+\pi) \tag{11}$$

$$R^r\equiv\frac{\Delta W_1^r}{W_0^r}=\frac{R-\pi}{1+\pi}\approx R-\pi \tag{12}$$

其中，$1+\pi\equiv(P_1^g/P_0^g)$。实际财富的比例变化就是你的实际收益R^r，它近似等于名义收益R减去通货膨胀率π。就连续复利收益而言，

$$\ln(W_1^r/W_0^r)\equiv R_c^r=\ln(1+R)-\ln(P_1^g/P_0^g)=R_c-\pi_c \tag{13}$$

这里的R_c等于（连续复利）名义收益，π_c为连续复利通货膨胀率。使用连续复利收益的好处在于，在$t=0$到$t=n$的时期中，实际收益的对数具有可加性：

$$\begin{aligned}R_c^r(0\to n)&=(R_{c1}-\pi_{c1})+(R_{c2}-\pi_{c2})+\cdots+(R_{cn}-\pi_{cn})\\&=R_{c1}^r+R_{c2}^r+\cdots+R_{cn}^r\end{aligned} \tag{14}$$

由上式可知，如果初始财富为W_0^r，那么在$t=n$时的实际财富水平为

$$W_n^r=W_0^r e^{R_c^r(0\to n)}=W_0^r e^{(R_{c1}^r+R_{c2}^r+\cdots+R_{cn}^r)}$$

一种替代表述是，如果使用比例变化收益，那么

$$W_n^r=W_0^r(1+R_1^r)(1+R_2^r)\cdots(1+R_n^r) \tag{15}$$

从$t=0$到$t=n$的年几何平均实际收益$\bar{R}_{r,g}$为

$$\begin{aligned}(1+\bar{R}_{r,g})&=\sqrt[n]{(1+R_1^r)(1+R_2^r)\cdots(1+R_n^r)}\\W_n^r&=W_0^r(1+\bar{R}_{r,g})^n\end{aligned}$$

国外投资

假设你正在考虑海外投资的问题。可以证明，用国内货币度量的名义收益等于外币收益（有时也被称为投资国当地货币收益）加上外币升值。进行海外投资，你要么从所持有的国外资产上要么从汇率的变化上获得收益（或遭受损失）。例如，考虑初始财富为W_0的一位英国居民，他将名义初始财富W_0（英镑）以S_0的汇率兑换成美元，并在美国进行投资，投资的名义（比例）收益为R^{us}。$t=1$时名义的英镑财富为

$$W_1=\frac{W_0(1+R^{us})S_1}{S_0} \tag{16}$$

因此，运用$S_1=S_0+\Delta S_1$就得到英国投资者海外投资的（比例）名义收益

$$\begin{aligned}R(\mathrm{UK}\to\mathrm{US})&\equiv(W_1/W_0)-1\\&=R^{us}+\Delta S_1/S_0+R^{us}(\Delta S_1/S_0)\\&\approx R^{us}+R^{FX}\end{aligned} \tag{17}$$

这里的$R^{FX}=\Delta S_1/S_0$为美元对英镑汇率的升值比率，我们假设$R^{us}(\Delta S_1/S_0)$可以忽略不计。显然，国外投资的名义收益为

名义收益（英国居民）= 当地货币收益（美国）+ 美元升值

就连续复利收益而言，方程的准确形式为：

$$R_c\ (\mathrm{UK}\to\mathrm{US}) \equiv \ln\ (W_1/W_0) = R_c^{\mathrm{us}} + \Delta s \tag{18}$$

其中，$R_c^{\mathrm{us}} \equiv \ln\ (1+R^{\mathrm{us}})$，$\Delta s \equiv \ln\ (S_1/S_0)$。现在，假设你关心的是国外投资的实际收益，即相对于国内商品购买力而言的收益。国外投资的实际收益恰好等于名义收益减去通货膨胀率。为了说明这一点，假设一位英国居民在美国进行投资，但所有的利润最后都用来消费英国商品。相对于英国商品的购买力而言，$t=1$ 时的实际财富为

$$W_1^r = \frac{(W_0^r P_0^g)\ (1+R^{\mathrm{us}})\ S_1}{P_1^g S_0} \tag{19}$$

于是，国外投资的连续复利（比例）实际收益为

$$R_c^r\ (\mathrm{UK}\to\mathrm{US}) \equiv \ln\ (W_1^r/W_0^r) = R_c^{\mathrm{us}} + \Delta s - \pi_c^{\mathrm{uk}} \tag{20}$$

$$R^r\ (\mathrm{UK}\to\mathrm{US}) \equiv \Delta W_1^r/\Delta W_0^r \approx R^{\mathrm{us}} + R^{\mathrm{FX}} - \pi^{\mathrm{uk}} \tag{21}$$

其中，$\Delta s = \ln\ (S_1/S_0)$。因此，相对于英国的购买力而言，英国居民对美国资产投资的实际收益 R^r（UK→US）为：

实际收益（英国居民）= 美国的名义“当地货币”收益 + 美元升值 − 英国的通货膨胀率

根据（20），注意如下事实是非常有趣的：如果

$$\pi_c^{\mathrm{uk}} - \pi_c^{\mathrm{us}} = \Delta s \tag{22}$$

那么英国居民国外投资的实际收益 R^r（UK→US）等于美国居民在美国投资的实际收益（$R_c^{\mathrm{us}} - \pi_c^{\mathrm{us}}$）。

正如我们在第 24 章中将要看到的那样，方程（22）就是相对购买力平价（PPP）条件。因此，如果相对购买力平价理论成立，那么国外投资的实际收益等于投资地的实际货币收益 $R_c^{\mathrm{us}} - \pi_c^{\mathrm{us}}$，汇率变化恰好被两个国家之间的通货膨胀率之差所抵消。因为相对购买力平价理论只在 5～10 年的时期中成立，所以短期国外投资的实际收益依赖于汇率变化。

1.2　折现现值（DPV）

假设 n 年内绝对安全投资的报价年利率为 r_n。每年计息一次的 A 美元现值在 n 年后的未来价值为

$$\mathrm{FV}_n = A\ (1+r_n)^n \text{美元} \tag{23}$$

因此，如果给你一个在 n 年后确定获得 FV_n 美元的机会，那么你愿意放弃今天的 A 美元。n 年后确定性支付 FV_n 美元在今天的价值为 A 美元。用更为专业的语言来说，FV_n 的折现现值（DPV，discounted present value）为

$$\mathrm{DPV} = \mathrm{FV}_n/\ (1+r_n)^n \tag{24}$$

现在，我们假设 1，2，3，…，n 年期的无风险利率为常数且都等于 r，假设利率期限结构是平坦的。无任何违约风险的现金流 FV_i（$i=1,\ 2,\ \cdots,\ n$）的 DPV 为

$$\mathrm{DPV} = \sum_{i=1}^{n} \mathrm{FV}_i/\ (1+r)^i \tag{25}$$

年金

如果未来支付在每一年都是常数（$\mathrm{FV}_i = C$ 美元）且第一笔支付是在第一年的年末进行的，那么我们拥有的就是普通年金（ordinary annuity）。这些支付的 DPV 为

$$\mathrm{DPV} = C \sum_{i=1}^{n} 1/\ (1+r)^i \tag{26}$$

利用几何级数的求和公式，我们可以将普通年金的 DPV 写为

$$\mathrm{DPV} = C \cdot A_{n,r} \tag{27}$$

其中，

$$A_{n,r} = (1/r)\ [1-1/(1+r)^n]$$

当 $n \to \infty$ 时，

$$\mathrm{DPV} \to C/r$$

称 $A_{n,r}$ 为年金因子（annuity factor），对 n 和 r 的各种取值而言，年金表给出了年金因子的相应数值。年金公式的一种特殊形式是，当 n 趋于无穷大时，$A_{n,r} = 1/r$ 且 $\mathrm{DPV} = C/r$。人们用这个公式来对被称为永久年金（perpetuity 或 console）的债券进行定价，这种年金的息票支付为 C 美元（但发行者永远不能赎回年金）。年金公式可用于诸如按揭、养老金和付息债券定价等涉及常数支付的相关计算（参见下文）。

实际投资项目

考虑诸如新建工厂之类的实际投资项目，这些项目有一系列的预期净收益（利润）FV_i。假设项目的资金成本为 KC 美元，项目从今天开始（即 $T=0$ 时）进行增值。那么企业家进行投资的条件是

$$\mathrm{DPV} \geqslant \mathrm{KC} \tag{28}$$

或等价地，净现值 NPV 必须满足

$$\mathrm{NPV} = \mathrm{DPV} - \mathrm{KC} \geqslant 0 \tag{29}$$

如果 NPV = 0，那么可以证明，投资项目的净收益（利润）恰好能偿还项目融资贷款的本金（KC 美元）和利息。如果 NPV > 0，那么在对这些贷款进行偿还之后还可以获得一笔剩余资金。

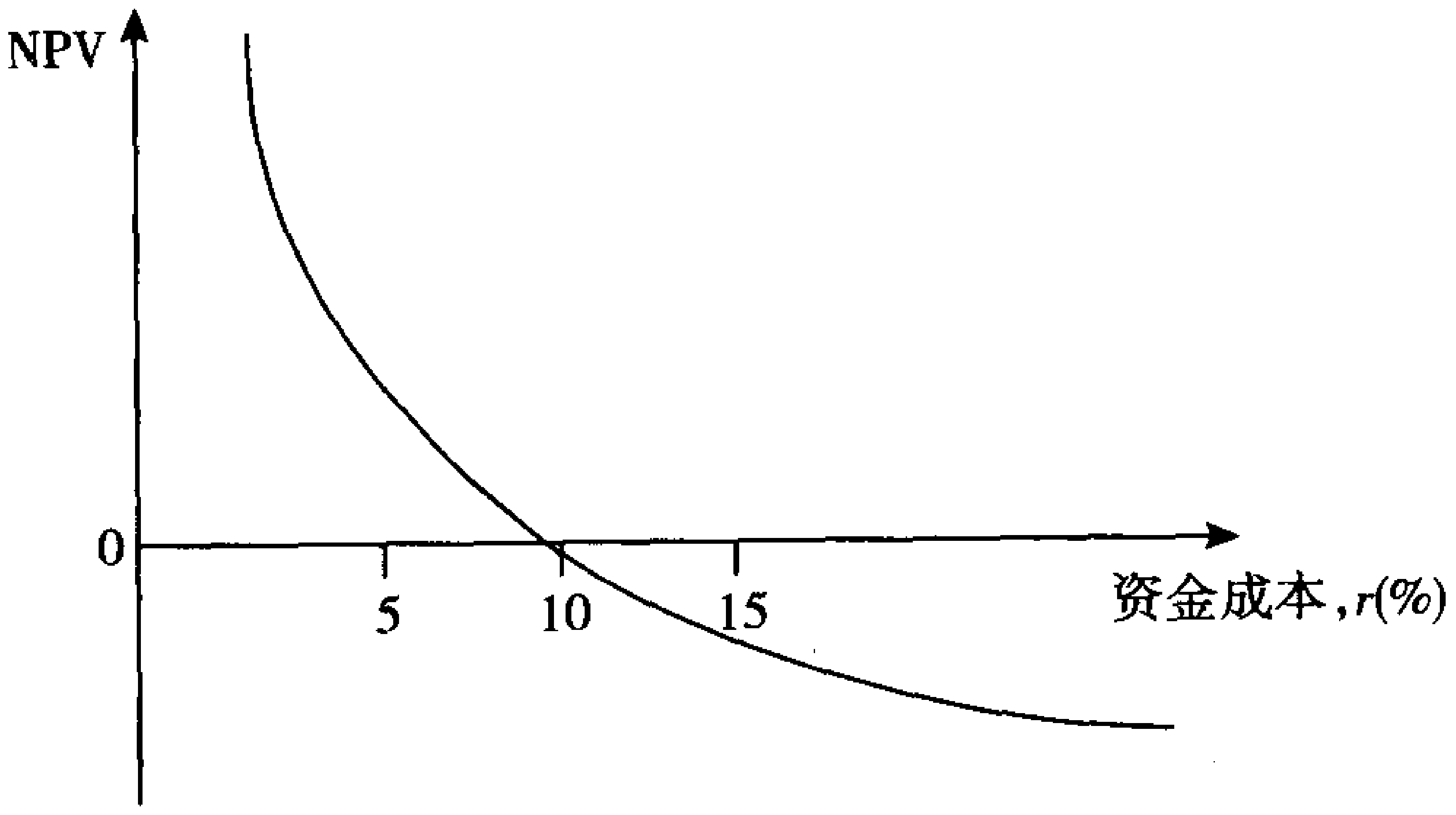


图 1　NPV 和贴现率

随着资金成本 r 的增加，对任意给定的项目利润流 FV_i 而言，NPV 将减小（图 1）。由 NPV = 0 就可以得出一个 r 值（图 1 中的 10%）。r 就是众所知周的投资项目内部收益率（IRR，internal rate of return）。在项目净收益流 FV_i 和资金成本给定的条件下，我们总能计算出项目的 IRR。IRR 就是满足

$$\mathrm{KC} = \sum_{i=1}^{n} \frac{\mathrm{FV}_i}{(1+y)^i} \tag{30}$$

的常数值 y。与 NPV 条件（28）等价的项目投资法则是：如果

$$\mathrm{IRR}\ (=y) \geqslant \text{借贷成本}\ (=r) \tag{31}$$

成立，那么就对项目进行投资。对 IRR 而言，通常存在一些技术问题（在实践中，没有任何问题的情形是非常幸运的）。首先，要使 IRR 的解有意义就必须假设所有 FV_i 都大于 0，因此 FV_i 不会改变符号。因为如果不是这样的话，IRR 的解就不止一个。其次，IRR 不能用于对两

个项目进行比较，因为它所给出的投资决策可能与NPV给出的投资决策不一致（参见Cuthbertson和Nitzsche（2001a））。

从本章开始，这些投资法则的使用贯穿全书，票据和债券收益率的计算以及经济中投资项目最优规模的确定等，都要用到这些投资法则。注意，在DPV的计算过程中，我们假设对未来收益FV_i进行折现的利率在所有时期中都为常数。假设“一年期货币”所承载的利率为r_1，两年期货币的成本为r_2，其余类似。因此，DPV为

$$\mathrm{DPV}=\frac{\mathrm{FV}_1}{1+r_1}+\frac{\mathrm{FV}_2}{(1+r_2)^2}+\cdots+\frac{\mathrm{FV}_n}{(1+r_n)^n}=\sum\delta_i\cdot FV_i \qquad (32)$$

其中，$\delta_i=1/(1+r_i)^i$。r_i就是众所周知的即期利率（spot rate），因为你的放款利率在0到1年的时期中为r_1，在0到2年的时期上为r_2，其余类似（用年复利利率进行表示）。在任何时点上，无违约资产的即期利率r_i与期限之间的关系就是众所周知的收益曲线（yield curve）。例如，如果$r_1<r_2<r_3$，其余类似，那么就称收益曲线是向上倾斜的收益曲线。短期利率变化和长期利率变化之间的关系是利率期限结构的一个研究主题。

DPV公式也可以表示成实际值的形式。在这种情形中，未来收益FV_i要用总商品价格指数来消除通货膨胀的影响，贴现因子要用实际利率来进行计算。

一般而言，实际投资项目不会没有任何风险，因为未来收益是不确定的。DPV计算中的这些不确定性可以用很多方法来进行处理。也许最简单的方法就是让贴现率δ_i由无风险即期利率r_i和风险溢价rp_i来产生：

$$\delta_i=(1+r_i+rp_i)^{-1} \qquad (33)$$

这也是我们即将采用的方法。方程（33）是一个恒等式，除非我们已经有风险溢价模型，否则无法使用这个公式。第3章将考察各种风险溢价模型。

股票

对股票直接应用DPV概念的困难在于，未来红利是不确定的，贴现因子也可能具有时变（time varying）特征。可以证明（见第4章），基础价值（fundamental value）V_t是未来红利DPV的数学期望：

$$V_t=E_t\left[\frac{D_{t+1}}{(1+q_1)}+\frac{D_{t+2}}{(1+q_1)(1+q_2)}+\cdots\right] \qquad (34)$$

这里的q_i是从$t+i-1$到$t+i$时期上的单期收益。如果那些信息充分的理性交易者之间买卖股票不会产生任何系统性获利机会，那么真实的股票市场价格P_t必须等于股票基础价值V_t。例如，如果$P_t<V_t$，那么投资者将购买那些价值被低估的股票，当P_t向V_t方向增加时，投资者就获得了资本收益。在有效市场上，这种获利机会瞬间就会消失。

显然，我们无法通过直接计算V_t的方式来考察V_t与P_t是否相等，因为预期红利（和贴现率）是无法观测的。然而在后面的章节中，我们将讨论那些能克服这个问题的方法，在$P_t=V_t$的意义下来考察股票市场是否有效。如果在DPV公式中增加一些简化假设（例如，预期未来红利以一个固定的速度增加，贴现率在每个时期上都为常数，即$q=R$），那么（34）就变成

$$V_0=D_0(1+g)/(R-g) \qquad (35)$$

这就是著名的Gordon增长模型。使用这个公式，我们能计算出股票的“公平”价值，将“公平”价值与市场报价P_0进行比较就可以考察股票是被高估了还是低估了。这些模型通常也被称为红利价值评估模型，我们在第10章对其进行讨论。

纯贴现债券和即期收益

考虑在纯贴现债券（零息债券）上的一笔投资，但不考虑实际投资项目上的投资。市场上通常将这些债券称为“零息债券”（zeros）。纯贴现债券有固定的赎回价格M，到期日是已知的，不支付任何息票。如果持有债券一直到期，那么债券的收益就由如下事实来确定：债券

的购买价格 P_t 低于赎回价格 M。就一年期债券而言，用下述公式计算债券收益率或利率是明智的：

$$r_{1t}=(M_1-P_{1t})/P_{1t} \tag{36}$$

这里的 r_{1t} 是用比例形式表示的收益率。然而，当我们根据 DPV 来看待这个问题时，结果就变成，一年期债券承诺在年末支付 M_1 来与今天的资本成本 P_{1t} 进行交换。因此，可以根据下式来计算债券的 IRRy_{1t}

$$P_{1t}=M_1(1+y_{1t}) \tag{37}$$

进行重新排列后，我们有 $y_{1t}=(M_1-P_{1t})/P_{1t}$，一年期即期收益率 r_{1t} 就简化为债券的 IRR。如果将上述原理应用到赎回价格为 M_2 的两年期债券上，那么债券的年（复利）收益率 r_{2t} 就是如下方程的解

$$P_{2t}=M_2/(1+r_{2t})^2 \tag{38}$$

这意味着

$$r_{2t}=(M_2/P_{2t})^{1/2}-1 \tag{39}$$

如果即期利率是连续复利利率，那么

$$P_{nt}=M_n e^{-r_{nt}n} \tag{40}$$

这里的 r_{nt} 是期限为 n 的债券在 t 时刻的连续复利利率。我们大体上已经看清楚了如何根据不同期限的纯贴现债券在 t 时刻的市场价格来计算 r 时刻的即期（复利）利率序列。

付息债券

付息（不可赎回）债券每隔一段时间（我们所取的时间间隔为 1 年）支付 C 美元的息票，在 n 年后债券到期时，债券固定的赎回价格为 M_n。对还有 n 年才到期的债券而言，当前的市场价格为 P_{nt}。问题是：如果持有债券到期，应该如何计算债券的收益率呢？

债券与今天延后 P_{nt} 美元资本而每年获得 C 美元收益（再加上赎回价格）的实际投资项目相类似。债券的内部收益率，我们称之为到期收益率 y_t，可以根据如下公式来进行计算

$$P_{nt}=C/(1+y_t)+C/(1+y_t)^2+\cdots+(C+M_n)/(1+y_t)^n \tag{41}$$

在未来偿付的 DPV 与当前市场价格相等的那些时点上，到期收益率为一常数，且等于贴现率。因为 P_{nt}、M_n 和 C 在市场中都是已知的常数，所以求解方程（41）就可以给出到期收益率 y_t 的市场报价。y_t 有一个下标"t"，因为随着市场价格的下降，根据"精算算术"算法，到期收益率会上升（反之，亦反之）。尽管到期收益率在市场和金融媒体中得到了广泛的应用，但将到期收益率作为持有债券一直到期的一种模棱两可的收益度量指标来使用存在一些理论/观念上的问题。后面部分将涉及这些问题。

债券在市场上通常是每 6 个月支付一次息票 C，从而（41）中的利率就是每 6 个月的每期利率。如果所计算的每期到期收益率为（比方说）6%，那么到期收益率的市场报价就很简单，即每年 12%（在美国，这就是著名的债券等价收益率）。

永久年金是一种初始发行者永远不能进行赎回的付息债券（即 $n\rightarrow\infty$）。如果每年的息票支付为美元，债券当前的市场价格为 $P_{\infty,t}$，那么由（41），永久年金的到期收益率为

$$y_{\infty,t}=C/P_{\infty,t} \tag{42}$$

由（42）立即就可以很明显看出，对很小的变化而言，永久年金价格变化的百分比等于到期收益率变化的百分比。扁平的收益或收益率或年收益率为 $y_n=(C/P_{nt})\times 100$，金融媒体所报道的就是这个收益率，但对债券收益和定价的分析而言，这不是一个很特别的理论问题。

尽管市场所报道的是复利利率（或收益率），我们还是经常发现，用连续复利即期利率/收益率来表示债券的价格要更为方便一些。如果连续复利即期收益率为 r_{nt}，那么可以将付息债券视为"零息债券"的一个组合，价格可以表示为（参见 Cuthbertson 和 Nitzsche（2001a））

$$P_{nt}=\sum_{k=1}^{n}C_k e^{-r_{kt}k}+M_n e^{-r_{nt}n}=\sum_{k=1}^{n}P_{kt}^{*}+P_{nt}^{*} \tag{43}$$

这里的 $P_k^* = C_k e^{-r_k k}$ 和 P_n^* 是在 $t+k$ 和 $t+n$ 时刻分别支付 C_k 和 M_n 的零息债券的价格。

持有期收益

与股票有关的许多实证研究都涉及单期持有期收益率（one - period holding return）H_{t+1}，定义如下

$$H_{t+1} = \frac{P_{t+1} - P_t}{P_t} + \frac{D_{t+1}}{P_t} \tag{44}$$

第一项是用比例表示的资本收益或损失（在单期上），第二项是红利收益（比例）。可以在事后对 H_{t+1} 进行计算，但是，从 t 时刻来看，P_{t+1} 和（或许）D_{t+1} 当然都是不确定的，投资者只能试着对它们进行预测。H_{t+1} 也满足如下关系式

$$1 + H_{t+i+1} = \frac{P_{t+i+1} + D_{t+i+1}}{P_{t+i}} \tag{45}$$

这里的 H_{t+i} 是从 $t+i$ 到 $t+i+1$ 的单期持有期收益率。因此，从事后来看，如果在股票上投资 A 美元（所有红利收益都要在股票上进行再投资），那么 n 期后 Y 美元的偿付等于

$$Y = A\ [1 + H_{t+1}]\ [1 + H_{t+2}]\ \cdots\ [1 + H_{t+n}] \tag{46}$$

连续复利持有期收益率（或“长期收益率”）定义为

$$h_{t+1} = \ln\ (P_{t+1}/P_t)\ = p_{t+1} - p_t \tag{47}$$

从 t 到 $t+n$ 的连续复利收益率定义为

$$h_{t+n} = p_{t+n} - p_t = h_t + h_{t+1} + \cdots + h_{t+n} \tag{48}$$

在整本书中，我们将说明预期单期收益率 H_{t+1} 是如何与 DPV 公式直接联系的。讨论股票市场是否有效的许多早期实证研究关注的都是单期收益率 H_{t+1} 是否可预测的问题。后来的实证研究关注的是股票价格是否等于未来红利的 DPV 值，最近的实证研究将两个分支的文献合在一起来进行讨论。

进行稍微修改后，我们可以针对任意资产定义单期持有期收益率。对期限为 n、息票支付为 C 的付息债券而言，我们有

$$H_{n,t+1} = \frac{P_{n-1,t+1} - P_{n,t}}{P_{n,t}} + \frac{C}{P_{n,t}} \tag{49}$$

也经常称它为（单期）持有期收益率（HPY）。注意，在 $t+1$ 时刻，n 期债券变为 $n-1$ 期债券。第一项是债券的资本收益，第二项是息票收益（或年收益）。对零息债券而言，$C=0$。在期限结构模型中，一般使用零息债券的连续复利收益，因此 h_{t+1} 为

$$h_{t+1} = p_{n-1,t} - p_{n,t} \tag{50}$$

我们经常使用相同类型的经济模型来解释股票和债券（及其他投机性资产）持有期收益率的运动，下一章以资本资产定价模型（CAPM）作为开始来进行我们的分析。

1.3　效用与无差异曲线

这一节简略地讨论效用的概念，目的只是为了便于读者能继续学习投资组合选择和随机贴现因子（SDF）模型方面的内容。经济学家频繁地建立投资组合模型，模型中的个体选择一组资产来使最终财富或投资组合收益或消费的函数达到最大。例如，当个体企图消费财富所能购买的产品或服务时，一定的财富水平等价于个体在某种程度上的满足。尽管他的财富可以翻倍，但他的满足程度却无法翻倍。另外一个例子是，如果某人每晚喝一瓶白酒，那么多消费一瓶白酒所带来的额外满足感就没有消费第一瓶白酒的满足感强。这就是边际效用递减假设。效用理论也可用于那些涉及不确定性（严格地说是“风险”）结果的决策。事实上，根据投资者效用函数的类型，可以将投资者分为“风险规避”、“风险偏好”和“风险中性”三种类型。

最后，我们也能对个体如何评估“效用”的问题进行考察，这个问题涉及不同的时点，也就是说，问题涉及多期或跨期模型中贴现效用的概念。

公平彩票

公平彩票（赌博）的定义是预期价值为零的彩票（例如，抛掷硬币，头面朝上（赢）就得到1美元，头面朝下（输）就损失1美元）。风险规避意味着个体不会接受这样的“公平”彩票，可以证明，风险规避也意味着个体的效用函数是关于财富的凹函数。考虑如下随机偿付 x：

$$x=\begin{cases}k_1, & \text{概率为 } p\\ k_2, & \text{概率为 } 1-p\end{cases} \tag{51}$$

公平彩票的预期价值必须等于零

$$E(x)=pk_1+(1-p)k_2=0 \tag{52}$$

这意味着 $k_1/k_2=-(1-p)/p$ 或 $p=-k_2/(k_1-k_2)$。就我们“抛掷硬币”的例子而言，$p=1/2$，$k_1=-k_2=1$ 美元。

预期效用

假设期末财富 W 是一随机变量，它有 n 个可能的取值：取 W_i 的概率为 p_i（$\sum_{i=1}^{n} p_i=1$）。个体在财富 W_i 上所获得的效用记为 $U(W_i)$，风险结果的预期效用为

$$E[U(W)]=\sum_{i=1}^{n} p_i U(W_i) \tag{53}$$

不确定性与风险

在效用函数上施加的第一个限制条件是，偏好总是多多益善，从而 $U'(W)>0$，$U'(W)=\partial U(W)/\partial W$。现在考虑一个简单的赌博：抛掷硬币，“头面”朝上得到16美元，头面朝下得到4美元。假设抛掷硬币是公平的，那么头面朝上的概率为 $p=1/2$，风险结果的预期货币价值为10美元：

$$EW=pW_H+(1-p)W_T=(1/2)\times16+(1/2)\times4=10\text{美元} \tag{54}$$

可以看出，当赌博的进入成本为10美元时，这个赌博就是一个公平赌博，因为 $E(x)=EW-c=0$。个体愿意为该赌博支付多少呢？这取决于个体对待风险的态度。如果个体愿意支付10美元来参与这个赌博，她也接受了这个公平赌博，那么她就属于风险中性类型。如果你厌恶风险结果，那么你愿意持有10美元而不是进行这样的公平赌博（预期价值为10美元），你属于风险规避类型。风险偏好者为参与该项赌博愿意支付的金额超过10美元。

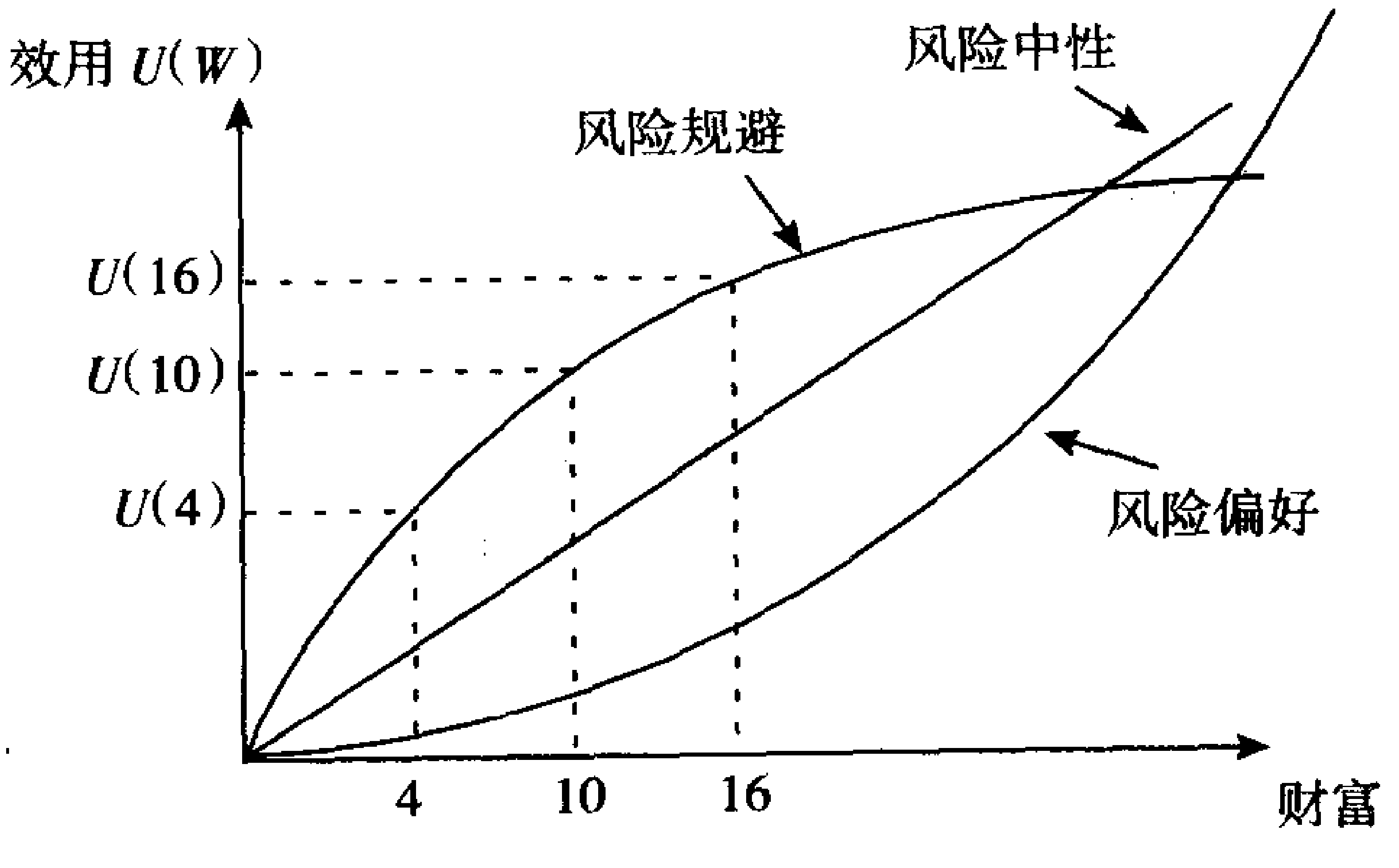


图 2　效用函数

风险规避意味着效用函数的二阶导数为负，$U''(W)<0$。为了看清这一点，请注意，为了使效用函数具有凹的形状，你持有 10 美元而放弃赌博的效用为 $U(10)$，这个效用必须大于参与赌博的预期效用：

$$U(10)>0.5U(16)+0.5U(4)$$

或

$$U(10)-U(4)>U(16)-U(10) \tag{55}$$

我们将这种类型的投资者称为“风险规避型投资者”，如图 2 所示。风险规避投资者效用函数的一个例子是 $U(W)=W^{1/2}$。注意，这个例子也适用于前面的公平赌博符号：如果 x 的风险结果为 $k_1=W_H-c=6$，$k_2=W_T-c=-6$，那么 $E(x)=0$。

也可以反过来叙述这个凹命题，凹性的意思就是投资者不愿意接受公平赌博。如果 z 是一个随机变量，$U(z)$ 是凹函数，那么由 Jensen 不等式有：

$$E\{U(z)\}<U\{E(z)\} \tag{56}$$

令 $z=W+x$，这里的 W 为初始财富。对一个公平赌博而言，$E(x)=0$，于是

$$E\{U(W+x)\}<U\{E(W+x)\}=U(W) \tag{57}$$

因此，你不会接受这个公平赌博。

很容易推导出如下结论：对风险偏好者而言，效用函数是财富的凸函数（如 $U=W^2$），而对风险中性者而言，赌博和确定性结果无差异，从而效用函数是线性函数（即 $U(W)=bW$，$b>0$）。因此，我们有：

对风险规避者而言，$U''(W)<0$；

对风险中性者而言，$U''(W)=0$；

对风险偏好者而言，$U''(W)>0$。

也可以这样说，风险规避者对财富的边际效用递减：如果初始财富水平越来越高，那么每一单位额外财富所增加的效用会越来越小（即 $U''(W)<0$）。风险规避程度由图 2 中效用函数的凹性来刻画，等价地说，也可以由 $U''(W)$ 绝对值的大小来刻画。注意，即使是很特殊的个体，风险规避程度也取决于初始财富和赌博规模。对上面的小额赌博而言，个体可能是风险中性的，愿意为参与赌博支付 10 美元。然而，如果赌博是“头面”朝上获得 1 百万美元，头面朝下获得 0 美元，那么赌博的预期价值就是 500 000 美元，但相同个体很有可能不愿意为规避赌博而支付 499 000 美元，即使这个赌博是他或她所喜欢的赌博（相同个体在大额赌博面前就变成了风险规避者）。当然，如果我们所说的个体是微软的比尔・盖茨，拥有一大笔初始财富，那么他可能愿意支付 500 000 美元来参与第二个赌博。

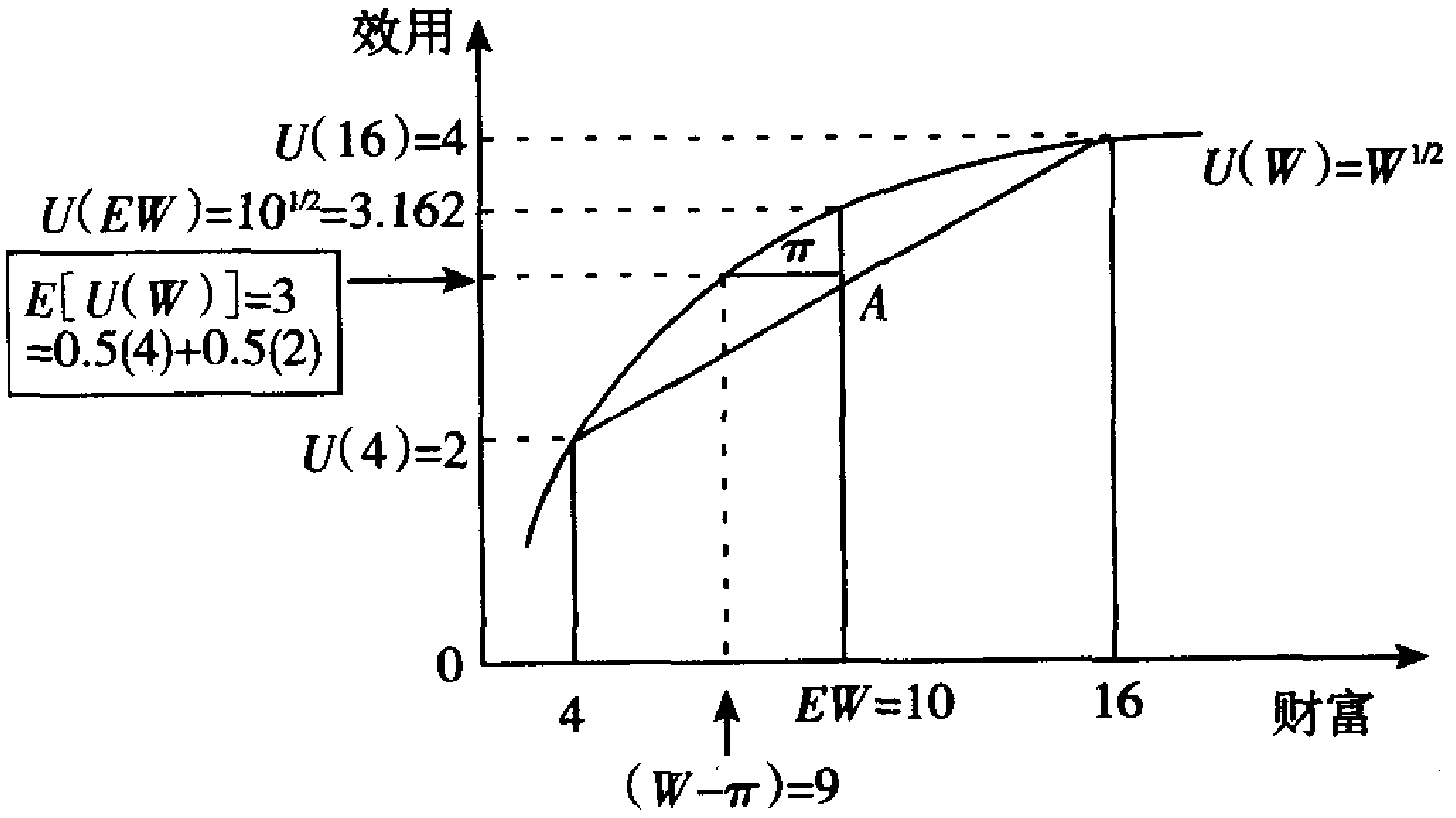


图 3　货币风险保费

风险规避意味着效用函数对风险赌博而言是一个凹函数。但是，我们怎样才能用货币而不是效用来对这种风险规避程度进行量化呢？答案就在图 3 中，这里的距离 π 就是你为规避这

个公平赌博所愿意支付的最大数目。如果你支付了 π，那么你将接受预期价值为确定性 10 美元的赌博，赌博的结果为（$10-\pi$）美元。假设风险规避型投资者的效用函数为 $U(W)=W^{1/2}$。赌博的预期效用为

$$E[U(W)]=0.5U(W_H)+0.5U(W_T)=0.5(16)^{1/2}+0.5(4)^{1/2}=3$$

注意，赌博的预期效用 $E[U(W)]$ 小于不参与赌博的确定性结果的效用 $U(EW)=10^{1/2}=3.162$。为避免参与赌博，风险规避型投资者愿意支付 $\pi=0.75$ 美元吗？如果他确实这样做，那么他的确定性效用为 $U=(10-0.75)^{1/2}=3.04$，大于赌博的预期效用 $E[U(W)]=3$，因此为了避免赌博他愿意支付0.75 美元。他愿意支付的最大保费 π 为多大呢？当从较低财富（$W-\pi$）处获得的确定性效用 $U(W-\pi)$ 恰好等于赌博的预期效用 $E[U(W)]=3$ 时，最大保费的情形就发生了：

$$U(W-\pi)=(10-\pi)^{1/2}=E[U(W)]=3$$

这个式子给出了你为避免参与赌博所愿意支付的最大数目，$\pi=1$ 美元。货币数量 π 就是众所周知的“风险保费”（risk premium），它是你为避免参与赌博所愿意支付的最高保费（注意，“风险溢价”（risk premium）在金融中还有另外一层含义，我们在后面将会遇到（即风险资产的预期收益超过无风险利率的部分））。公平赌博增加或减少 $x=6$ 美元就得到了点 A 处的预期效用。如果你的财富减少到（$W-\pi$），那么效用水平为 $U(W-\pi)$。因此，风险保费 π 就定义为

$$U(W-\pi)=E\{U(W+x)\} \tag{58}$$

这里的 W 为初始财富。为了看清 π 是如何与效用函数的凹性产生联系的，对（58）在 $x=0$（即概率密度集中在均值零的附近）和 $\pi=0$ 处进行 Taylor 级数逼近，

$$\begin{aligned}U(W-\pi)&\approx U(W)-\pi U'(W)\\&=E\{U(W+x)\}\\&\approx E\{U(W)+xU'(W)+\frac{1}{2}x^2U''(W)\}\\&=U(W)+\frac{1}{2}\sigma_x^2U''(W)\end{aligned} \tag{59}$$

因为 $E(x)=0$，所以我们要求 $U(W+x)$ 的展开式中有三项。根据（59），风险保费为

$$\pi=-\frac{1}{2}\sigma_x^2\frac{U''(W)}{U'(W)}=\frac{1}{2}\sigma_x^2R_A(W) \tag{60}$$

这里的 $R_A(W)=-U''(W)/U'(W)$ 是 Arrow（1970）－Pratt（1964）绝对（局部）风险规避度量指标。风险规避度量指标具有“局部性”，因为它是初始财富的函数。

因为 σ_x^2 和 $U'(W)$ 都为正，所以 $U''(W)<0$ 意味着 π 为正。注意，你为避免公平赌博所愿意支付的数量取决于结果的风险大小 σ_x^2，也取决于 $U'(W)$ 和 $U''(W)$。例如，你可能非常厌恶风险（$-U''(W)$ 很大），但你不愿意支付一个高的风险保费 π，同时你又很穷，那么 $U'(W)$ 就会很高。事实上，普遍使用的两个风险规避度量指标是：

$$R_A(W)=-U''(W)/U'(W) \tag{61}$$

$$R_R(W)=R_A(W)W \tag{62}$$

$R_A(W)$ 为 Arrow－Pratt（局部）绝对风险规避度量指标，$R_A(W)$ 越大，风险规避程度越高。$R_R(W)$ 是相对风险规避度量指标。当财富在初始（“局部”）财富水平附近发生变化时，投资者的风险偏好会发生怎样的变化呢，R_A 和 R_W 两个指标都可以对此进行度量。

不同数学函数所导出的风险规避有着不同的含义。例如，函数 $U(W)=\ln(W)$ 表明绝对风险规避是递减的，而相对风险规避却是常数（见后面）。现在，我们列出在资产定价和投资组合领域中经常使用的一些“标准”效用函数。

幂效用函数（常相对风险规避效用函数）

假设初始（安全）财富为 W_0，效用函数相对于起点具有如下性质：

$$U(W)/U(W_0)=f(W/W_0)$$

从而效用只对财富的相对差异产生反应，我们称这种效用函数为相对风险规避型效用函数。幂效用函数就满足这个条件，效用函数对 W/W_0 的反应在这里是常数，因此幂效用函数等价的术语是常相对风险规避（CRRA）效用函数：

$$\begin{aligned} U(W)&=\frac{W^{(1-\gamma)}}{1-\gamma}\quad \gamma>0,\ \gamma\neq 1\\ U'(W)&=W^{-\gamma}\\ U''(W)&=-\gamma W^{-\gamma-1}\\ R_A(W)&=\gamma/W\\ R_R(W)&=\gamma\ (\text{常数})\end{aligned}\tag{63}$$

因为 $\ln[U'(W)]=-\gamma\ln W$，所以 γ 也是效用函数相对于财富的边际效用替代弹性。

对数效用函数

在（63）中，令 $\gamma\rightarrow 1$，可以证明幂效用函数的极限情形就是对数效用函数：

$$U(W)=\ln(W)\ \text{且}\ R_R(W)=1\tag{64}$$

它有着漂亮而又简单的直观性质，当你的财富从 W 变成 W^2 时，你的满足程度（效用）也会翻倍。

二次效用函数

$$\begin{aligned} U(W)&=W-\frac{b}{2}W^2\quad b>0\\ U'(W)&=1-bW\\ U''(W)&=-b\\ R_A(W)&=b/(1-bW)\\ R_R(W)&=bW/(1-bW)\end{aligned}\tag{65}$$

因为 $U'(W)$ 必须为正，所以只有当 $W<1/b$ 时二次效用函数才有定义，这就是众所周知的“极乐点”（bliss point）。边际效用是财富的线性函数，这在有些时候是一个很有用的性质。注意，R_R 和 R_A 都不是常数，都是财富的函数。

负指数效用函数（常绝对风险规避效用函数）

假设初始（安全）财富水平为 W_0，效用函数相对于起点具有如下性质：

$$U(W)/U(W_0)=f(W-W_0)$$

从而效用只对财富的绝对差异产生反应，我们称这种效用函数为绝对风险规避型效用函数。满足这个要求的唯一（可接受）效用函数就是（负）指数效用函数，效用函数对 $W-W_0$ 的反应为常数，因此这个效用函数的等价术语是常绝对风险规避（CARA）效用函数：

$$\begin{aligned} U(W)&=a-be^{-cW}\quad c>0\\ R_A(W)&=c\\ R_R(W)&=cW\end{aligned}\tag{66}$$

可以证明，负指数效用函数加上资产收益正态分布的假设允许人们将预期效用最大化问题简化为如下问题：最大化投资组合预期收益 ER_p 和风险的线性函数，风险（确切地说）用方差 σ_p^2 来表示。于是，最大化上述 CARA 效用函数 $E[U(W)]$ 就等价于最大化

$$ER_p-\frac{c}{2}\sigma_p^2\tag{67}$$

这里的 c 是常绝对风险规避系数。方程（67）只取决于投资组合收益的均值和方差，由此就有了均值方差权衡这一术语。然而，读者将会注意到，就一般情况而言，最大化 $E[U(W)]$ 并

不能简化为如下问题：最大化问题的目标函数只取决于 ER_p 和 σ_p^2 的一般函数（见附录）。只有效用函数为负指数效用函数时，预期效用最大化问题才能简化为对一个线性函数进行最大化的问题。一些投资组合模型一开始就假设投资者只关心均值方差的最大化，因此它们也就放弃了与特殊效用函数有关的任何直接联系。

HARA（双曲型绝对风险规避效用函数）

$$U(W)=\frac{1-\gamma}{\gamma}\left(\frac{\alpha W}{1-\gamma}+\beta\right)^{\gamma} \tag{68}$$

$$R_A(W)=\left(\frac{W}{1-\gamma}+\frac{\beta}{\alpha}\right)^{-1} \tag{69}$$

$$R_A(W)>0,\ \gamma>1,\ \beta>0$$

限制条件为 $\gamma\neq 1$，$[\alpha W/(1-\gamma)]+\beta>0$，$\alpha>0$。另外，如果 $\gamma=-\infty$，那么 $\beta=1$。我们对 HARA（双曲型绝对风险规避效用函数）感兴趣的原因是，它包含了常绝对风险规避效用函数（$\beta=1$，$\gamma=-\infty$）、常相对风险规避效用函数（$\gamma<1$，$\beta=0$）和二次效用函数（$\gamma=2$）三种类型，文献中使用的效用函数都是它的特殊情形。

1.4 资产需求

我们经常想知道，在投资者的投资组合中，股票、债券和其他资产的最优需求由哪些因素决定。答案取决于如何设定最大化问题和投资者所面临的约束条件，这并不让人感到奇怪。我们在这里集中讨论单期问题，问题的解相对而言要简单一些，后续章节将涉及更为复杂的情形。

均值方差优化模型

最简单的模型就是假设投资者只关心投资组合的单期预期收益和投资组合收益的标准差（风险）。投资组合由两种资产构成：收益为 R 的单一风险资产，令 α 为该资产在初始财富 W_0 中的持有比例；收益为 r 的无风险资产，$(1-\alpha)$ 为其持有比例。预算约束为（劳动收入为零）为

$$W_1=(\alpha W_0)(1+R)+[(1-\alpha)W_0](1+r)$$

因此，投资组合的收益和标准差分别为

$$R_p\equiv\frac{W_1}{W_0}-1=\alpha(R-r)+r$$

$$\sigma_p=\alpha\sigma_R$$

这里的 σ_R 是模型中唯一随机变量 R 的标准差。假设投资者最大化

$$\max_{\alpha}\theta=ER_p-\frac{c}{2}\sigma_p^2$$

这里的 $c>0$ 是风险规避程度的度量指标（更为准确地说，在投资组合预期收益与投资组合收益的方差之间进行权衡）。优化问题的一阶条件是

$$ER-r-\alpha c\sigma_R=0$$

从而风险资产的最优份额与财富无关：

$$\alpha^*=\frac{ER-r}{c\sigma_R}$$

因此，风险资产的绝对（美元）数量 $A_0=\alpha^* W_0$ 与初始财富成比例，与风险资产的超额收益正相关，与风险规避程度和风险资产的波动率负相关。无风险资产的份额可以简单地表示为 $(1-\alpha^*)\equiv A_{0f}/W_0$。上述模型就是 Tobin（1956）资产需求的均值方差模型，模型能得到简单

闭型解（closed form solution）的原因在于，最大值函数关于 α 是二次函数（因为 $\sigma_p^2 = \alpha^2\sigma_R^2$）。如果将已知的非随机劳动收入包含在预算约束中，那么模型的解不会发生变化。理论文献有时也使用单期模型，因为最大值函数是预期收益的线性函数，这种线性性提供了易于处理的解析解。

很容易将均值方差方法扩展到 n 种风险资产 $R = (R_1, R_2, \cdots, R_n)'$ 的情形，最大值函数为

$$\max_{\alpha} \theta = \alpha'(R - r \cdot e) + r - \frac{c}{2}\alpha'\Omega\alpha$$

其中，$\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)'$，$e$ 是数值全为 1 的 $n \times 1$ 列向量，Ω 是收益的 $n \times n$ 方差协方差矩阵。一阶条件（FOC）为

$$\alpha^* = (c\Omega)^{-1}(ER - r \cdot e)$$

无风险资产的份额为 $\alpha_f^* = 1 - \sum_{i=1}^{n} \alpha_i^*$。就两种风险资产的情形而言，

$$\Omega^{-1} = (\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21})^{-1}\begin{pmatrix} \sigma_{22} & -\sigma_{21} \\ -\sigma_{12} & \sigma_{11} \end{pmatrix}$$

因此，预期收益（$ER_1 - r$）和（$ER_2 - r$）的相对权重取决于收益方差协方差矩阵的每一个元素。二阶条件保证了 $\partial\alpha_i^*/\partial ER_i > 0$（$i = 1$ 或 2）。

负指数效用函数

值得注意的是，在一般情形中，不可能通过对任意特殊效用函数 $U(W_1)$（这种效用函数只取决于最终财富）进行二阶 Taylor 级数展开的方法来得到最大值 θ。进行二阶 Taylor 级数展开通常会导致如下非线性函数 $EU(W) = U(ER_p) + \frac{1}{2}\sigma_p^2 U''(ER_p)$，而均值方差方法得到的是预期收益和方差的线性函数。然而，在某些（非常特殊的）情形中，我们能将最大值 θ 与一个特殊的效用函数直接联系起来，即

$$\max_{\alpha} E[U(W)] = -E\{\exp(-bW_1)\} = -E\{\exp(-bW_0(1 + R_p))\}$$

$$\text{约束条件：} R_p \equiv \frac{W_1}{W_0} - 1 = \alpha'(R - r \cdot e) + r$$

这里的 b 是常绝对风险规避系数。因此，效用函数必须是负指数函数（期末财富为 W_1），正如下面将要看到的那样，资产收益也必须服从多维正态分布。如果随机变量服从正态分布，那么 $z = \exp(x)$ 服从对数正态分布。z 的数学期望为

$$Ez = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

在我们的例子中，$\mu \equiv R_p$，$\sigma^2 \equiv \mathrm{Var}(R_p)$。最大值函数是关于指数部分的单调函数，因此，$\max E[U(W)]$ 等价于

$$\max_{\alpha} E[U(W_1)] = \alpha'(ER - r \cdot e) - \frac{1}{2}bW_0\alpha'\Omega\alpha$$

我们扔掉了非随机项 $\exp(-bW_0)$。现在，最大值函数与投资组合的收益和方差线性相关。解的一阶条件为

$$\alpha^* = (bW_0\Omega)^{-1}(ER - r \cdot e)$$

解的形式与均值方差情形中解的形式相同。如果 $c = bW_0$，那么两者等价。然而，要注意的是，由负指数函数所导出的资产需求函数意味着，在风险资产上投资的绝对美元数 $A = \alpha^* W_0$ 独立于初始财富。因此，如果个体在下一期获得额外财富，那么他会将所有的额外财富都投资于无风险资产（这个结果有些令人难以置信）。

二次效用函数

在这一节里，我们将勾画如下问题的轮廓：当效用函数是二次效用函数时如何推导资产份

额。数学推导相当杂乱（但不难），因此我们要对记号进行简化（与 Cerny（2004）一样）。假设只存在一种风险资产和一种无风险资产。预算约束为

$$W=\tilde{\alpha}W_0R^*+(1-\tilde{\alpha})W_0R_f^*+\gamma$$

其中，$\tilde{\alpha}=A/W_0$ 是风险资产的份额，γ 是已知的劳动收入，$R^*=1+R$ 是风险资产的总收益。无风险资产的份额为 $\tilde{\alpha}_f=1-\tilde{\alpha}$。进行重新排列之后，预算约束变成

$$W=W_{\text{safe}}(1+\alpha X)$$

其中，

$$W_{\text{safe}}=R_f^*W_0+\gamma$$
$$X=R-R_f$$
$$\alpha=\tilde{\alpha}W_0/W_{\text{safe}}$$

因此，α 仅仅是对 $\tilde{\alpha}$ 进行尺度变换后的版本。当财富的“幸福水平”（bliss level）W_{bliss} 固定时，效用函数为二次效用函数：

$$U(W)=-\frac{1}{2}(W-W_{\text{bliss}})^2=-\frac{1}{2}(W^2-2WW_{\text{bliss}}+W_{\text{bliss}}^2)$$

假设投资者总是处于极乐点之下。由上式很容易看出，$E[U(W)]$ 取决于 α、α^2、EX 和 $E(X^2)$，从而一阶条件是 α 的线性函数。另外，$E(X^2)=\text{Var}(X)-(EX)^2$，从而最优的 α 将只取决于风险资产的期望超额收益 EX 和风险资产收益的方差（但这种关系不是线性关系）。将预算约束代入到 $E[U(W)]$ 并对一阶条件关于 α 进行求解，我们就得到（要经过冗长的代数运算）

$$\alpha^*=q_k\frac{EX}{E(X^2)}$$
$$q_k=2k(1-k)/2k^2$$
$$k=W_{\text{safe}}/W_{\text{bliss}}$$

注意，在 α^* 方程中，风险规避度量指标没有直接出现，但风险度量指标隐含在效用函数的平方项中，从而它是解 α^* 的一个尺度变换。（我们在下面也将看到，幂效用函数的解可以简化为 $\gamma=-1$ 时二次效用函数的解。）

为了不超过极乐点，我们要求 $k\ll 1$；为了简化代数运算，取 $k=1/2$ 来使 $q_k=1$（方程 3.53，Cerny（2004），p. 68）。因此

$$\alpha^*=\frac{EX}{E(X^2)}=\frac{\mu_k}{\sigma_x^2+\mu_x^2}=\frac{1}{\mu_x(1+1/SR_x^2)}$$

$SR_x=\mu_x/\sigma_x$ 是众所周知的夏普比。在本书中，夏普比是每单位风险收益（回报风险比率）的度量指标。最优值 α（和 $\tilde{\alpha}$）在这里与夏普比直接相关。可以证明，二次效用函数情形（$W(\alpha^*)<W_{\text{bliss}}$）中的 α^* 也是对夏普比进行最大化时所给出的值（在多种风险资产的情形中，也可以对这个结论进行一般性推广（参见 Cerny（2004），第 4 章））。

因此，我们可以将“选择投资组合（即 $\alpha=(\alpha_1, \alpha_2, \cdots, \alpha_n)$）来最大化夏普比的问题”与“最大化预期二次效用（$W(\alpha^*)<W_{\text{bliss}}$）的问题”直接联系起来（尽管这种情形可能出现：如果最优的 α 满足 $W(\alpha^*)>W_{\text{bliss}}$，那么夏普比准则无效）。对其他效用函数而言，我们无法建立（基础）夏普比与效用之间的这种联系。然而，人们在实践中经常使用夏普比来对各种风险组合进行排序，但不试图将决策与任何特殊效用函数联系起来（正如我们在后面章节中要看到的那样）。此外，如果我们能将上述情形推广到将“基础”夏普比与常 CARA 效用函数联系起来，那么我们就称这个夏普比为 Hodges 比率（见 Cerny（2004），Hodges（1998））。

幂效用函数

对绝大多数效用函数而言，得到资产份额的闭型解是不可能的，只能使用数值技巧。假设

最终财富由一种风险资产和一种无风险资产组成，效用函数依赖于（单期）最终财富，我们将说明幂效用函数所对应的数值结果。预算约束为

$$W(\tilde{\alpha})=\tilde{\alpha}W_0R^*+(1-\tilde{\alpha})W_0R_f^*+y=W_{\text{safe}}(1+\alpha X)$$

假设存在如下简单设定：$R_u^*=1.20$，$R_D^*=0.90$。因此，风险资产只有两种结果：分别以 1/2 的概率向上或向下（即 20% 或 -10%）运动。令 $\gamma=5$，$r=0.03$，$W_0=1$ 美元，$y=$ 200 000美元。那么幂效用函数所对应的优化问题为

$$\begin{aligned}\max_{\alpha}\theta&=E\left[\frac{W^{1-\gamma}}{1-\gamma}\right]\\&=\frac{1}{2}\frac{W_u^{1-\gamma}}{1-\gamma}+\frac{1}{2}\frac{W_D^{1-\gamma}}{1-\gamma}\\&=W_{\text{safe}}^{1-\gamma}\left\{\frac{1}{2}\frac{1+\alpha X_u}{1-\gamma}+\frac{1}{2}\frac{(1+\alpha X_u)^{1-\gamma}}{1-\gamma}\right\}\end{aligned}$$

除 α 外，θ 中所有参数都是已知的。事实上，在这种情形中得到解析解是可能的。令 $\partial\theta/\partial\alpha=0$，根据所得到的方程来求解 α^*，因此 $\tilde{\alpha}=\alpha^*W_{\text{safe}}/W_0=0.3323$（见 Cerny（2004），p60）。另外一种方法的思想是：任何使用数值方法来实现最大化的投资者都能直接得出问题的解。将所得到的 α^* 代入上述方程就可以得到最优点处的预期效用：

$$\theta^*=E[U(W(\alpha^*))]=-1.110\,397\ (10^{-25})$$

等价的确定性财富水平 W_{cert} 可以通过下式来进行计算

$$U(W_{\text{cert}})=E\{U[W(\alpha^*)]\}\text{，即}\ \frac{W_{\text{cert}}^{1-\gamma}}{1-\gamma}=-1.110\,397\ (10^{-25})$$

当 $\gamma=5$ 时，$W_{\text{cert}}=1\,224\,942$ 美元。$W_{\text{safe}}=R_f^*W_0+y=1\,220\,000$ 美元，因此，与只持有无风险资产的情形相比，投资者的福利增加了 4 942 美元。

很容易将上述讨论扩展到如下情形：投资组合仍然只包含一种风险资产，但风险资产的超额收益 X 有 m 种可能（“状态”），第 i 个结果发生的概率为 p_i，对所有状态求和就可以得到最大值 θ，

$$\theta=\max_{\alpha}\sum_{i=1}^{m}p_iU(W_i)$$

其中，$U(W_i)=W_i^{1-\gamma}/(1-\gamma)$，$W_i=W_{\text{safe}}(1+\alpha X_i)$。再一次出现的情形是，$\theta$ 中唯一的未知参数就是风险资产的份额 α（或 $\tilde{\alpha}$），最优的 α 也可以通过数值优化方法来获得（或者，如果你有很多纸张的话，那么你也可以“用手进行计算”）。

现在，让我们更为大胆一些，假设选择的是一组风险资产 $\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n)'$。但是，就此刻而言，假设每一个超额收益的自然状态都只有 $m=4$ 种，超额收益的联合概率分布为

$$\begin{aligned}&X^{(i)}=(X_1^{(i)},\cdots,X_n^{(i)}) &&(n\times1)\\&W_i=W_{\text{safe}}(1+\alpha'X^{(i)}) &&(i=1-4)\\&p_i=(p_1,p_2,p_3,p_4) &&(1\times4)\end{aligned}$$

W_i 是一个指标，n 维向量 $X^{(i)}$ 每一种结果的联合概率为 p_i。对 4 维向量 $X^{(i)}$ 而言，只有 4 种可能结果，概率分别为 p_i，从而 W_i 和 $U(W_i)$ 也只有 4 种结果。因此，θ 应在 4 个状态上进行求和，很容易将它的值计算出来，但 θ 现在取决于 n 个 α_i 的值。一阶条件为 $\partial\theta/\partial\alpha_i=0$（$i=1,2,\cdots,n$），就一般情形而言，这些非线性方程无法求得解析解，必须使用最优化方法才能得到使 θ 最大的 α_i^*（$i=1,2,\cdots,n$）。

连续分布

假设我们拥有 n 种资产但收益为 R 的分布连续，为表述方便起见，假设 $n\times1$ 向量 $X=R-r\cdot e$（这里的 e 是数值全为 1 的向量 $n\times1$）服从多维正态分布，条件密度函数为 $f(x\mid\Lambda)$，

这里的 Λ 是 $t-1$ 时刻及之前的所有信息构成的集合。我们有

$$W=W_{\text{safe}}(1+\alpha'X)$$

$$U(W)=W^{1-\gamma}/(1-\gamma)$$

于是，

$$\theta=E\{U[W(\alpha)]\}=\frac{W_{\text{safe}}^{1-\gamma}}{1-\gamma}\cdot\int_{-\infty}^{\infty}(1+\alpha'x)^{1-\gamma}f(x\mid\Lambda)\,\mathrm{d}x$$

为了说明问题，假设 $n=2$，且 $X\mid\Lambda\sim N(\mu,\Omega)$，于是

$$f(x)=\frac{1}{(2\pi)^{n}\mid\det\Omega\mid}\exp\left[-\frac{1}{2}(x-\mu)\Omega^{-1}(x-\mu)\right]$$

假设条件均值和协方差矩阵都是“已知的”（即由统计学家根据历史数据来进行估计），中括号中的项是关于 x 的一个纯量函数（scalar function）。在最简单的情形中，x_{it}关于时间独立同分布，但在时刻 t 存在交叉相关性，

$$x_i=\mu+\varepsilon_i \qquad \Omega=E(\varepsilon'\varepsilon)$$

条件矩和无条件矩相等。（多重）积分可以用解析方法来进行求解，人们通常却使用数值方法来进行计算。事实上，最大化者选择 $\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 的各种可能值来计算 $\theta(\alpha)$，选择 $\alpha=\alpha^*$ 来得到最大值 $\theta(\alpha^*)$。Cerny（2004）第 4 章给出了最优化过程的一个清晰描述，包括一些有用的 GAUSS 程序。

风险规避与投资组合选择

当投资者只考虑一个时期并追求 $U(W)$ 的最大化时，就风险资产需求对初始财富变化的反应而言，我们几乎没有什么可以多说。为简单起见，假设只存在一种风险资产。我们不加证明地叙述一些结果，但对任何凹效用函数而言，这些结果都是建立在一阶条件分析基础上的：

$$E[U'(W)(R-R_f)]=0$$

其中，$W=W_0(1+R_f)+A(R-R_f)$，W_0 是初始财富，A 是在风险资产上的投资数量（(W_0-A) 是在无风险资产的投资数量）。根据一阶条件很容易看出第一个结论。如果 $ER=R_f$，那么 $A=0$ 满足一阶条件，$W=W_0(1+R_f)$ 是非随机的，因此

$$E[U'(W)(R-R_f)]=U'(W)E(R-R_f)=0$$

因此，风险规避个体不愿意接受公平赌博，即 $ER-R_f=0$，因为后者意味着 $A=0$。对凹效用函数而言，我们有如下结果：

（ⅰ）如果 $ER>R_f$，那么 $A>0$，投资者持有的风险资产数量为正，他愿意接受预期收益为正的小额赌博。

（ⅱ）递减的绝对风险规避系数（即 $\partial R_A/\partial W_0<0$，这里的 R_A 为绝对风险规避系数）意味着 $\partial A/\partial W_0>0$，也就是说，如果初始财富越高，那么个体投资者愿意将更多的“美元”投资到风险资产上去（反之，亦反之）。

（ⅲ）如果相对风险规避系数 R_R 是关于财富的递减函数（即 $\partial R_R/\partial W_0<0$），那么 $(\partial A/A_0)/(\partial W/W_0)>1$，因此随着财富的增加，个体在风险资产上的投资比例会增加。当个体投资者的相对风险规避系数是财富的递增函数时，就会得到相反的结果。

上述结果是对前面结论的一个补充。就常绝对风险规避（CARA）效用函数（如负指数效用函数）而言，$\partial A/\partial W_0=0$；就常相对风险规避系数效用函数而言，$\partial A/A=\partial W/W$，最优的资产份额仍为常数。

1.5 无差异曲线与跨期效用函数

尽管这是某些受限环境下的唯一情形，我们还是假设图 2 中风险规避者的效用函数可以唯

一地表示为投资组合的预期收益和收益方差的函数。期末财富 W 与预期收益为 ER_p 的资产组合之间的联系是 $W = (1 + ER_p) W_0$，这里的 W_0 为初始财富。然而，我们现在要假设效用函数可以表示为

$$U = U(ER_p, \sigma_p^2) \tag{70}$$

$$U_1 > 0, \ U_2 < 0, \ U_{11}, \ U_{22} < 0$$

一阶偏导数符号（U_1，U_2）意味着预期收益可以使效用增加，而更多的“风险”会使效用降低。二阶偏导数意味着，额外预期“收益”的边际效用递减，而额外风险的边际效用却递增。上述效用函数的无差异曲线如图 4 所示。

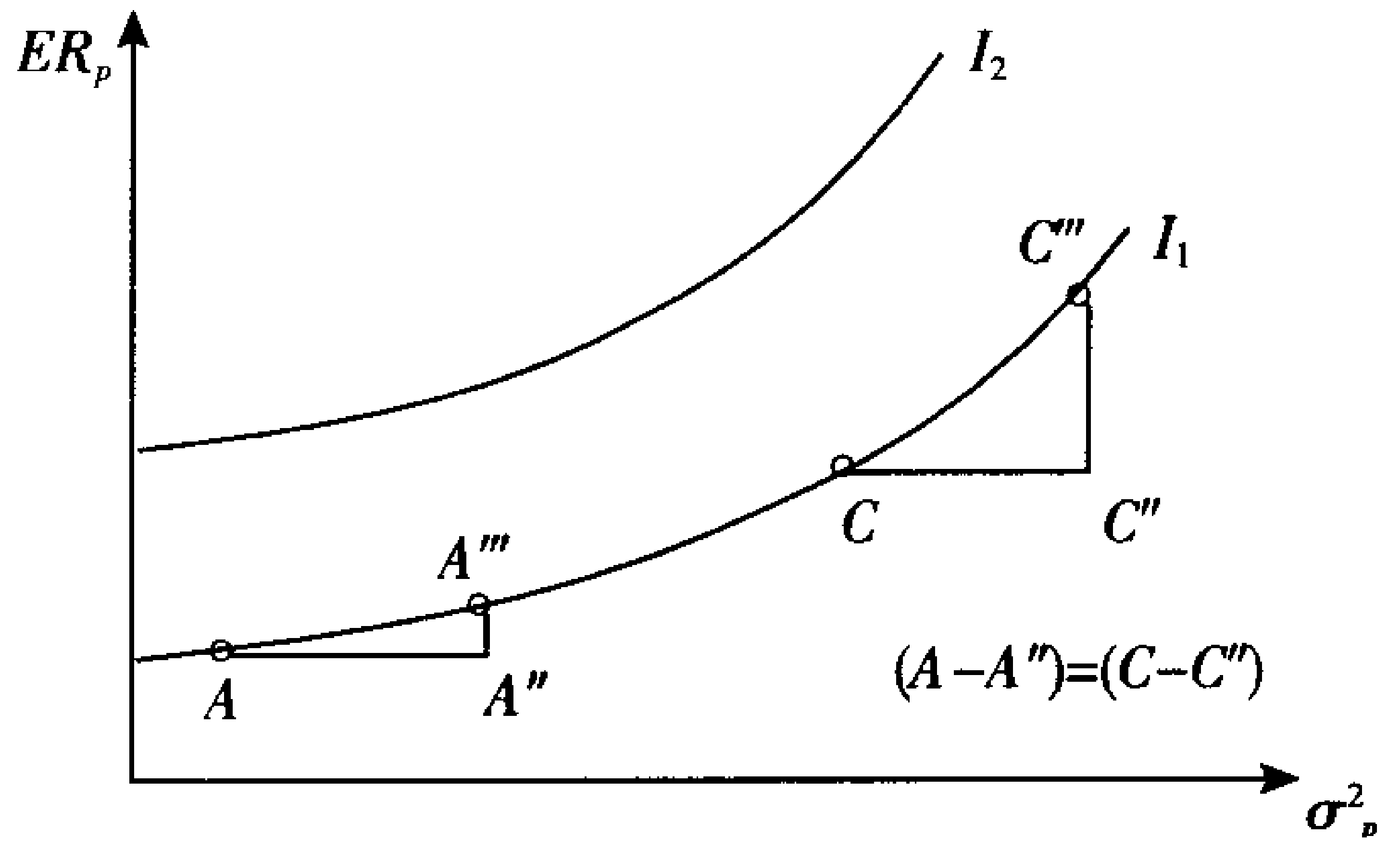


图 4　风险—收益：无差异曲线

在无差异曲线 I_1 上的 A 点，如果个体为了维持 A 点的满足程度（效用），那么作为更高风险（$A - A''$）的回报，他会要求一个更高的预期收益（$A''' - A''$）：无差异曲线在风险－收益空间上的斜率为正。无差异曲线凸向“风险轴”，这意味着在更高的风险水平上，比如说 C 点，个体会要求比 A 点更高的预期收益（$C''' - C'' > A''' - A''$），个体属于“风险规避”类型。在简单的均值方差模型中对投资组合选择问题进行分析时，我们将使用风险收益空间的无差异曲线。

跨期效用

个体行为的许多经济模型都假设投资者只从消费品中获得效用。在任何时点上，效用是消费的递增函数，且边际效用递减，

$$U = U(C_t) \qquad U'(C_t) > 0, \ U''(C_t) < 0 \tag{71}$$

因此，效用函数与图 2 中“风险规避者”的斜率相同（用 C 代替 W）。唯一的新增问题是，我们应如何处理不同时点上的消费增加问题。这种跨期终身效用函数最为一般的形式是

$$U_N = U(C_t, C_{t+1}, C_{t+2}, \cdots, C_{t+N}) \tag{72}$$

然而，为了使这里的数学易于处理，通常要对 U 的形式施加一些限制，最为常见的形式是主观贴现率为常数（$0 < \theta < 1$）的可加性效用函数：

$$U_N = U(C_t) + \theta U(C_{t+1}) + \theta^2 U(C_{t+2}) + \cdots + \theta^N U(C_{t+N}) \tag{73}$$

终身效用函数能用有限值 N 来进行截断，如果 $N \to \infty$，那么就称模型为世代交叠模型，因为个体消费会遗留给下一代。

（73）中所使用的贴现率取决于当前消费和未来消费之间“感觉”的比较。如果定义 $\theta = 1/(1+d)$，那么 d 就是众所周知的主观时间偏好率（subjective rate time preference）。它是个体将 $t+j$ 时刻的效用与 $t+j+1$ 时刻的效用进行交换并使终身效用保持不变的比率。（73）中的可加性意味着，第 t 年额外消费的边际效用与其他任何年份额外消费的边际效用（进行合适的折现）独立。

就两期情形而言，我们能画出一些简单效用函数（如 $U = C_0^{\alpha_1} C_0^{\alpha_2}$，$0 < \alpha_1$，$\alpha_2 < 1$）的无差异曲线，如图 5 所示。$A$ 点位于比 B 点更高的无差异曲线上，因为个体在 A 点与在 B 点有着相

同的 1 期消费水平 C_1，但他在 A 点的零期消费 C_0 要更高一些。在 H 点，如果将 C_0 减少 x_0 个单位，那么对那些保持终身效用水平不变的个体而言，他必须用 1 期的额外消费 y_0 来进行补偿，因此他在 H 点和 E 点上的效用是无差异的。边际效用递减是因为，在 F 点，如果他放弃 x_0 单位消费 C_0，那么他会要求 y_1（$>y_0$）单位 C_1 来对此进行补偿。这是因为，他在 F 点的初始消费 C_0 比 H 点低，从而他所放弃的每一单位 C_0 的相对价值都更大一些，所要求的补偿（C_1）也更高一些。

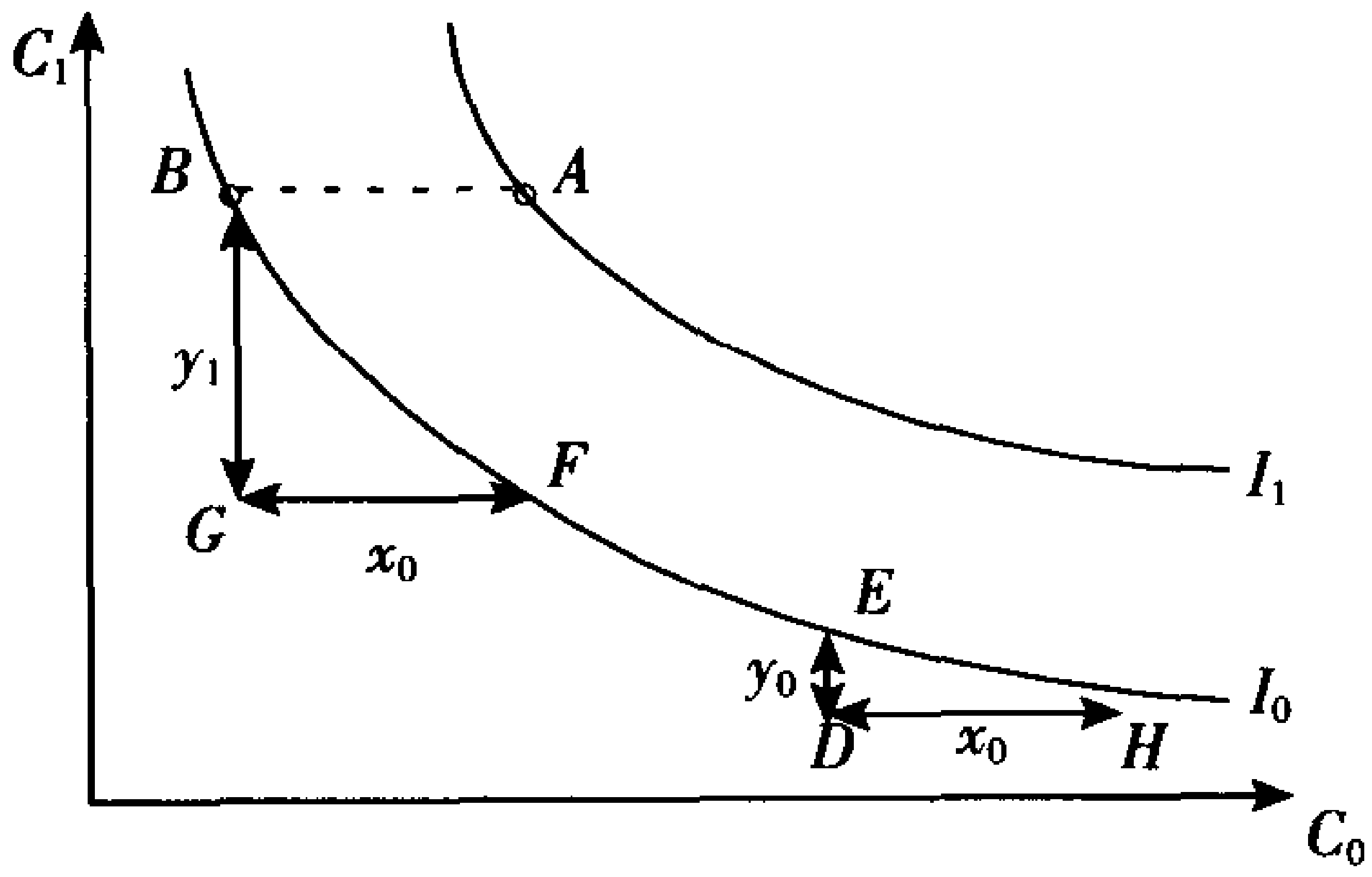


图 5　跨期消费：无差异曲线

下一节将用图 5 中的跨期无差异曲线来讨论确定状态下的投资决策问题、不确定性状态下投资组合选择的 CCAPM 模型和均衡资产收益问题。

1.6　投资决策与最优消费

在未来收益确定的条件下，我们的投资决策规则意味着，管理者要么按照 NPV 要么按照 IRR 来对实际投资项目进行排序。投资项目会被采纳直到最后一个被采纳项目的 NPV 等于零，或等价地说，IRR 等于无风险利率 r。在这些情形中，被采纳的边际（最后）投资项目所获得的净收益（利润）恰好能偿还贷款利息和本金。就整个经济而言，进行实际投资需要牺牲当前消费和产出。$t=0$ 时用于生产新机器或提高劳动技能的投入，如劳动技能、工作小时和机器等，只会增加未来各期的产出和消费。生产者决策所导致的消费变化（即减少今天的消费增加明天的消费）可能与个体消费者理想的消费变化不一致。例如，高水平的实际投资将极大地减少当前消费资源，对那些偏好今天消费而不是明天消费的消费者而言，这可能不是他们所愿意看到的。

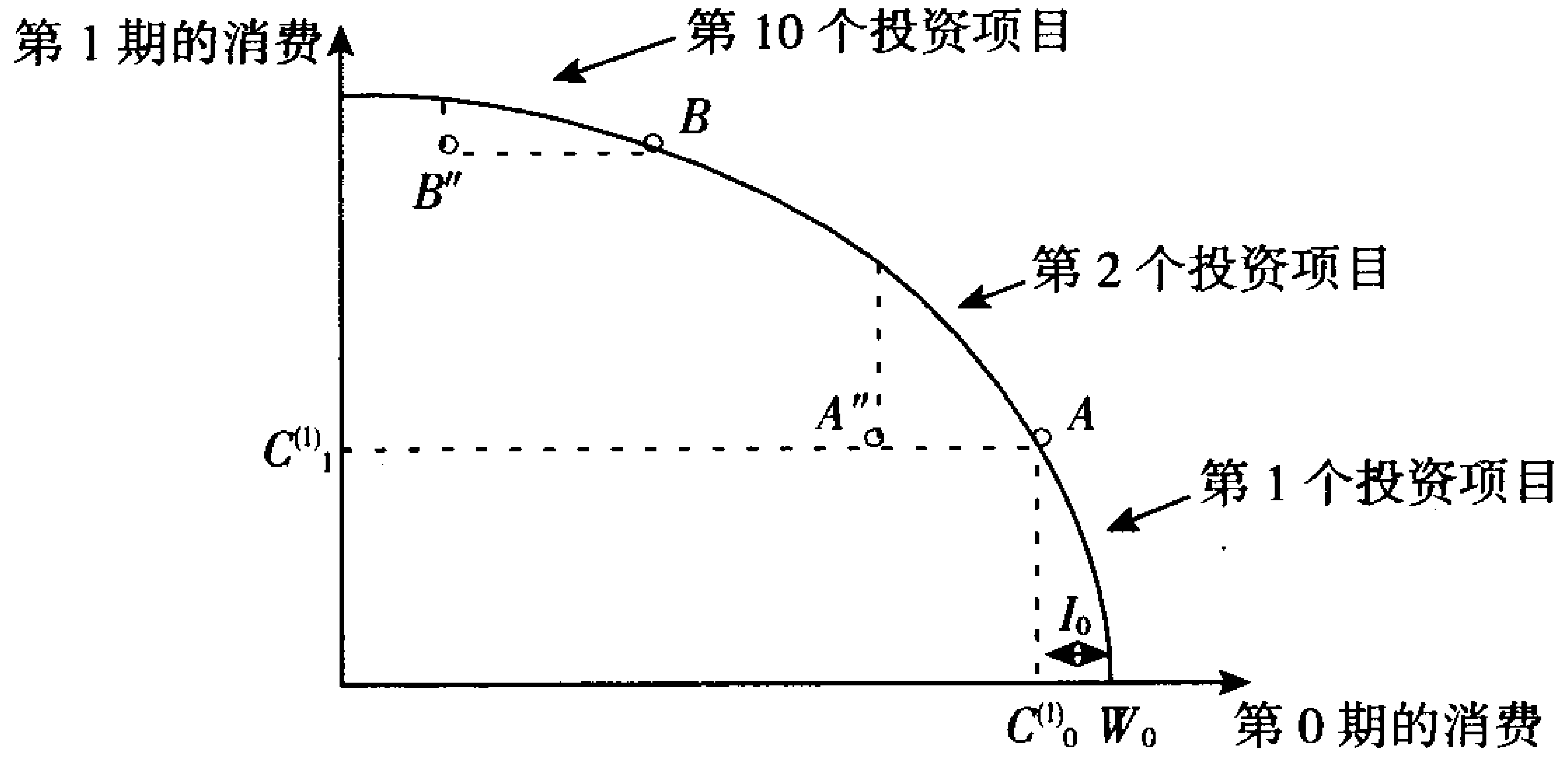
金融市场如何通过便利借贷来保证企业家在最优投资水平上进行投资（即能在未来生产出更多的消费品）并允许个体根据偏好在时间上匀滑消费？为了选择最优的实际投资水平，企业家必须知道个体消费者的偏好吗？作为股东的消费者，怎样才能保证企业经理做出“正确”的实际投资决策？我们能假设金融市场（即股票市场）将资金配置到最有效的投资项目上去吗？

“金融”与实际投资决策之间的交互作用问题是金融系统最为核心的问题，问题的完整回答涉及一些复杂问题。然而，如果考虑一个简单的两期投资决策模型，那么我们就能获得一些有用的直觉，模型中所有结果的实际值（即假设通货膨胀率为零）都是确定的（即无风险）。我们将看到，分离原理（separation principle）在这些假设下是适用的。如果管理者忽略个体偏好，简单地进行项目投资直到 NPV $=0$ 或 IRR $=r$，也就是说，他们追求公司价值的最大化，那么在给定的资本市场上，投资决策允许每个消费者选择合意的消费水平，即决策能使个体福利达到最大。因此，这是一个两阶段决策过程或决策分离过程；到目前为止，我们仍然允许消费者根据偏好在时间上分配消费来追求福利最大化。在第一个步骤中，企业家不管消费者的偏好

来决定最优投资水平。在第二个步骤中，消费者在资本市场进行借或贷来对不同时间上的消费进行重新安排，从而满足个体偏好。在解释分离原理时，我们首先涉及生产决策，其次是消费决策，最后将两者合在一个完整的模型中。

所有产出要么用于消费，要么用于实际投资。企业家在 $t=0$ 时的初始财富为 W_0。企业家根据 NPV 按照递减的方式对项目进行排序，并以无风险利率 r 作为贴现因子。通过节约消费 $C_0^{(1)}$，他获得了第一个投资项目所需要的资源 $I_0=W_0-C_0^1$。这个投资项目具有最高的 NPV（或 IRR），在 $t=1$ 时刻的消费产出为 $C_1^{(1)}$（$C_1^{(1)}>C_0^{(1)}$）（见图 6）。投资项目的 IRR 为（用消费来表示）

$$1+\mathrm{IRR}^{(1)}=C_1^{(1)}/C_0^{(1)} \tag{74}$$



注：$(A-A''=B-B'')$

图 6　生产可能性曲线

当他将更多的初始财富投资到那些较低 NPV 的项目上去时，IRR（C_1/C_0）就会下降，这会导致如图 6 所示形状的生产可能性曲线。第一个投资项目也是生产能力最强的投资项目，这个项目的 NPV 为

$$\mathrm{NPV}^{(1)}=C_1^{(1)}/(1+r)-I_0>0 \tag{75}$$

且

$$\mathrm{IRR}^{(1)}=C_1^{(1)}/C_1^{(0)}>r \tag{76}$$

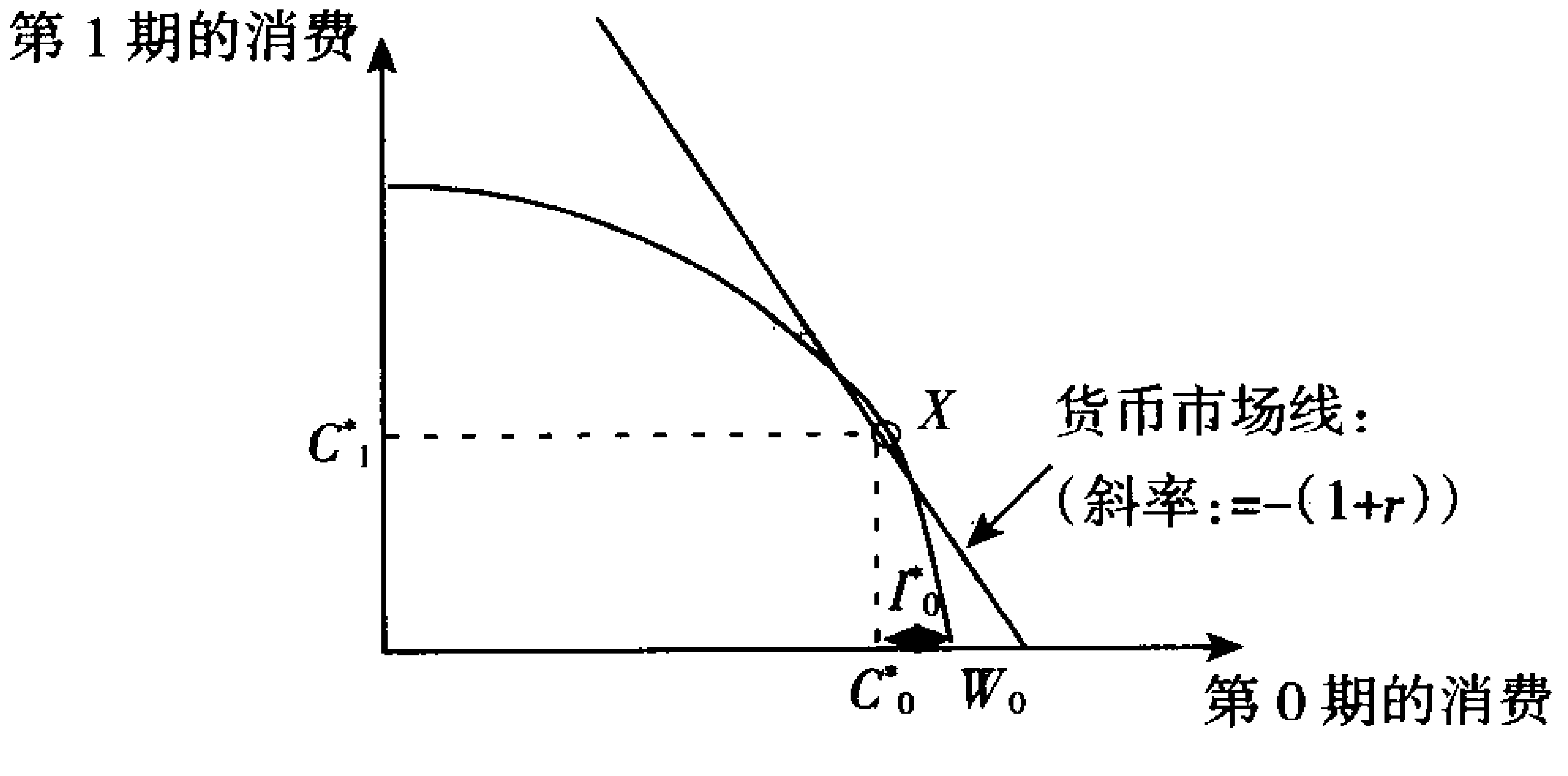


图 7　货币市场线

现在，我们转为讨论融资问题。在资本市场上，任意两个消费流 C_0 和 C_1 的现值（PV）为

$$\mathrm{PV}=C_0+C_1/(1+r) \tag{77a}$$

因此

$$C_1=\mathrm{PV}(1+r)-(1+r)C_0 \tag{77b}$$

对给定的 PV 值而言，上式给出的是图 7 中斜率为 $-(1+r)$ 的直线。我们称上述方程为“货

币市场线”（money market line），因为它代表的是贷款或借款的收益率。如果你今天贷出 C_0 的金额，那么你在明天将收到 $C_1 = (1+r)\ C_0$。

初始财富为 W_0 的企业家在实物资产上将继续投资直到第 n 个项目的 IRR 恰好等于无风险利率

$$\mathrm{IRR}^{(n)} = r$$

这种情形发生在（C_0^*，C_1^*）上。因此，追求公司价值（净现值）最大化的投资策略所涉及的投资金额为

$$I_0^* = W_0 - C_0^*$$

当前消费为 C_0^*，$t=1$ 时的消费为 C_1^*（图 7）。在 X 右边的任意时点上，投资机会曲线的斜率（$=$IRR）大于市场利率（$=r$），在 X 左边任何时点上，情形刚好相反。然而，生产决策的最优消费水平（C_0^*，C_1^*）可能与根据个体消费者偏好所得出的最优水平不一致。现在，我们离开生产决策话题，完全转到消费决策的讨论上去。

假设消费者的收入在两个时期中都会增值，收入流的现值为 PV。（两期之后）将收入完全耗费掉的消费可能性由（77a）给出。假设消费者的终身效用（满足）取决于 C_0 和 C_1

$$U = U\ (C_0,\ C_1)$$

C_0 和 C_1 的边际效用递减（即 $\partial U/\partial C > 0$，$\partial^2 U/\partial C^2 < 0$,）。无差异曲线如图 8 所示。如果消费者的目的在于维持初始的效用水平，那么为了放弃一单位 C_0，他必须获得额外的 C_1 单位补偿。消费者希望在预算约束范围内选择 C_0 和 C_1 来最大化终身效用。给定禀赋 PV，他在两期中的最优消费为（C_0^{**}，C_1^{**}）（图 8）。一般而言，能产出消费（C_0^*，C_1^*）的最优生产或实际投资计划不等于消费者的最优消费（C_0^{**}，C_1^{**}）。然而，资本市场的存在性保证了消费者的最优点是可以到达的。为了看清这一点，考虑图 9。

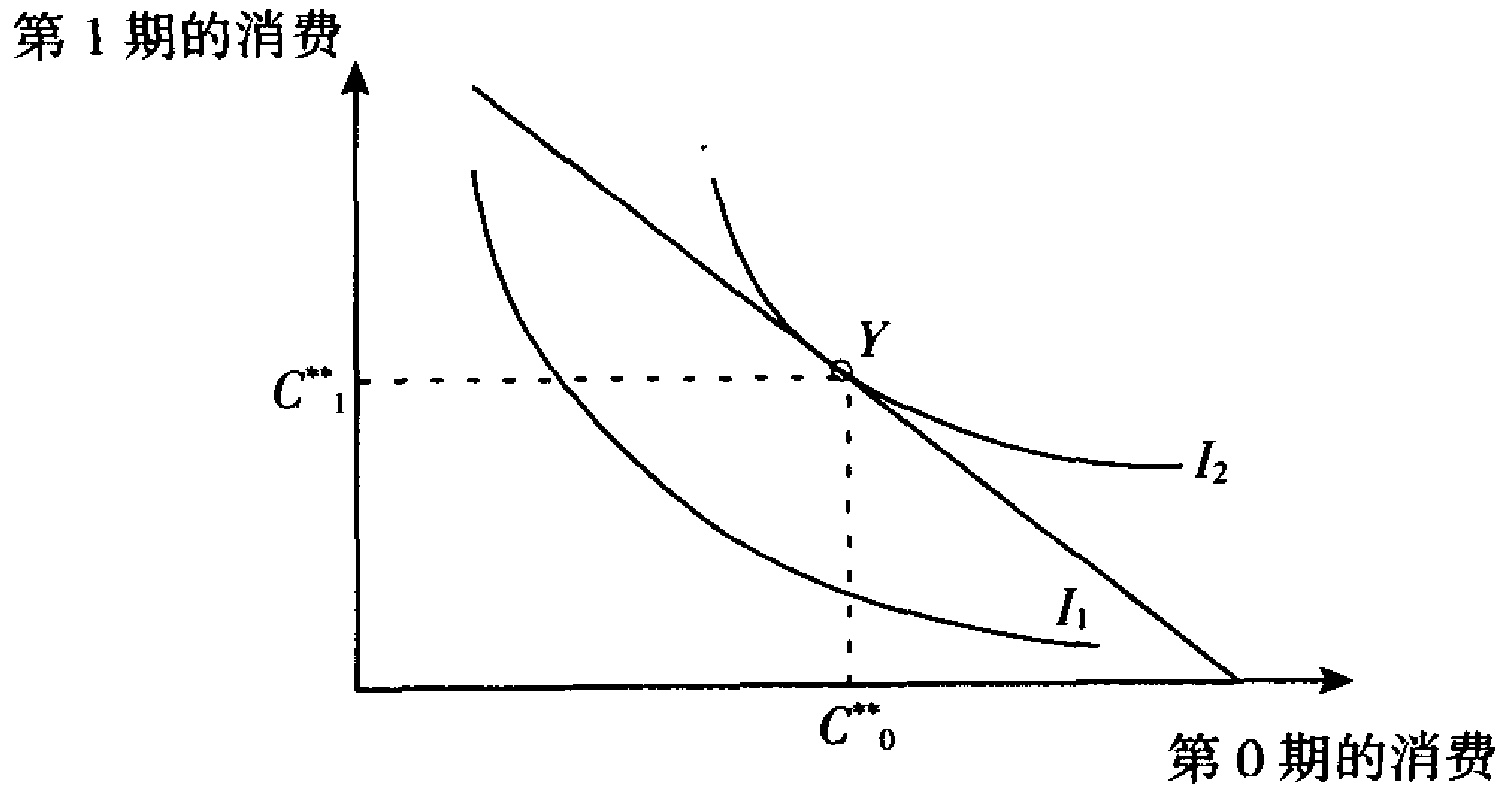


图 8　消费者的最大化问题

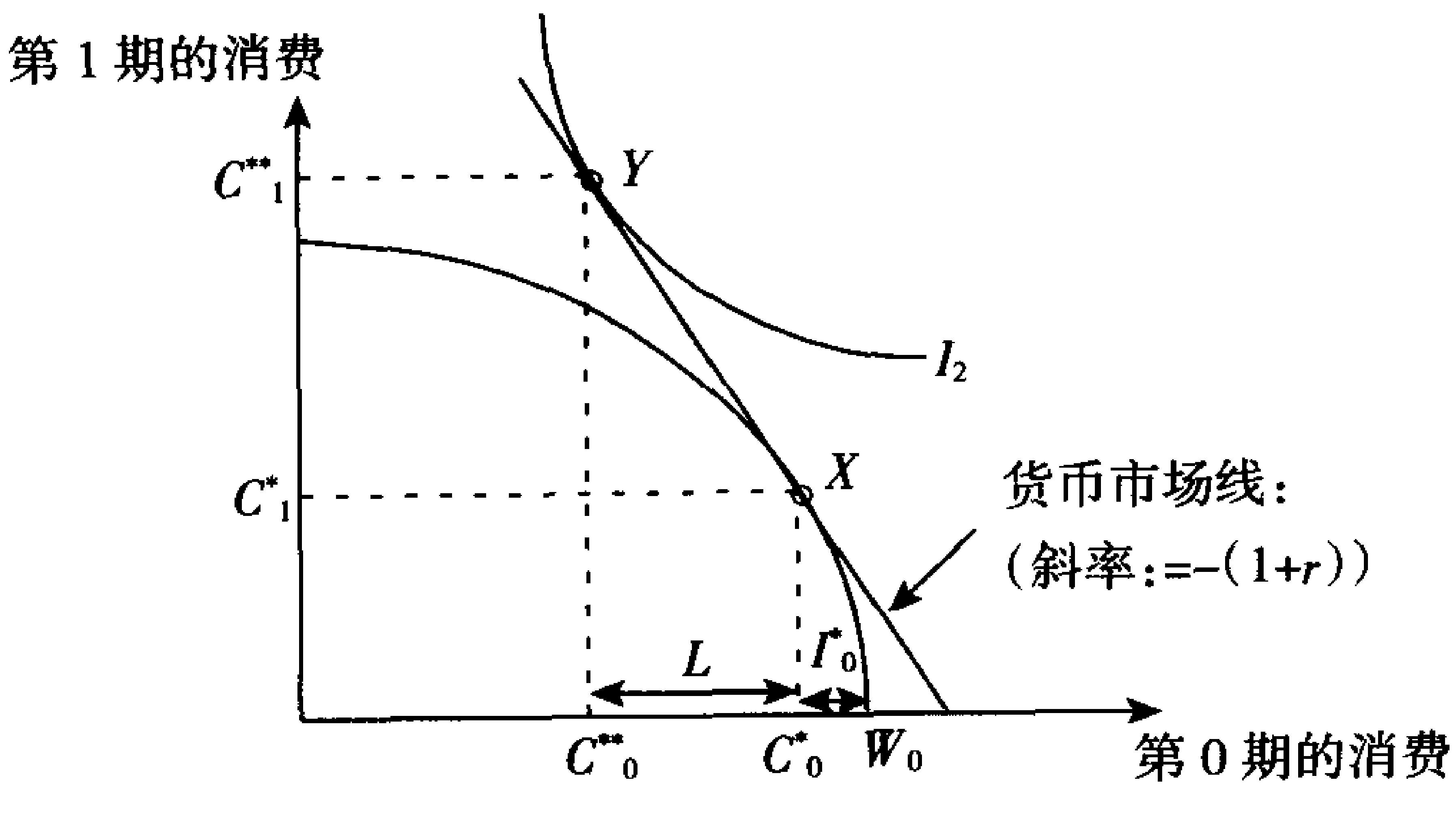


图 9　有资本市场时的最大化问题

企业家在追求企业价值最大化的过程中已经生产出消费（C_0^*，C_1^*），如图 9 所示。可以将消费视为企业所有者所支付的收入（红利），“现金流”的现值为 PV^*

$$PV^* = C_0^* + C_1^* / (1+r) \tag{78}$$

当然，“收入”是企业所有者支付给个体消费者的，消费者可以将 PV^* 的金额在任意消费组合之间进行“转换”，只要满足

$$PV^* = C_0 + C_1 / (1+r) \tag{79}$$

就可以。给定 PV^* 和图 9 中的无差异曲线 I_2，他能以无风险利率 r 在资本市场上进行借或贷，从而实现效用最大化的组合（C_0^{**}，C_1^{**}）。

因此，投资决策与融资决策（借和贷）之间是可以分离的。最优借贷决策独立于实际投资决策。如果企业家和消费者是同一个人，那么分离原理仍然成立。投资者（我们现在就这样称呼）首先决策初始财富 W_0 中多少用于实际资产投资，这一决策与他的（主观）偏好和感觉无关。第一阶段的决策是在对投资项目 IRR 与无风险利率进行比较基础上的一种客观计算。第二阶段的决策涉及在资本市场上借贷多少来“匀滑”不同时点上的消费使之到达合意的水平。第二阶段的决策是建立在偏好基础上的，要在今天消费与明天（更多）的消费之间进行边际权衡。

在本书余下的部分中，有很大一部分内容要考虑的问题是，在风险环境下如何进行融资决策。股东如何保证经理按照股东的利益来进行决策，即如何保证经理追求公司价值的最大化？在一些公司控制机制（如兼并，接管等）下，这样的问题就会产生。本书并不直接讨论公司控制机制。我们只考虑，市场价格是否为资源配置（即实际投资）提供了正确的信号，但不详细讨论在这些市场信号基础上的公司内部激励结构问题：这是公司金融的委托代理问题。

确定性环境下的上述结果与风险环境下即将发展的结果是平行的。

（ⅰ）在风险环境下，所应用的分离原理有一点细微差异。当投资者在市场上可交易风险资产（如股票、债券等）的投资组合中进行选择时，无论投资者对风险与收益的偏好如何，他所持有的风险资产份额与其他投资者都相同。在进行第一阶段决策后，每一个投资者需要考虑在货币市场上以无风险利率借或贷多少的问题（就是在这一点上，投资者的偏好对风险资产与无风险资产之间的分离存在一定的影响）。这个分离原理是最优投资组合选择的均值方差模型和均衡资产收益 CAPM 模型的基础。

（ⅱ）在个体当前与未来消费的主观边际替代率（即 $(\partial C_1/\partial C_0)_u$）等于 $-(1+r)$ 的那些时点上，投资者可以在货币市场上以无风险利率借贷最优数量的资金，这里的 r 表示货币的“价格”或机会成本。在不确定性环境下，类似的条件也成立，即个体在预期收益与风险之间的主观权衡等于风险的市场价格。

1.7　小结

我们已经发展了对金融市场上行为进行分析的一些基础工具。所讨论的主题之间存在一些细微差别，我们没有时间对此进行详细叙述，后面的章节将对这些忽略进行修正。主要结论有：

* 在一般情形中，市场参与者的报价利率是“单利”年利率，但这些利率总可以转换为有效年（复利）利率或连续复利利率。

* DPV 和 IRR 的概念能用于实际投资项目的分析，计算债券和股票的公平价格。

* 资产需求和资产价格的理论模型经常将效用函数当作客观函数来使用。效用函数及无差异曲线能用来描述风险规避者、风险偏好者和风险中性者的行为。

* 在确定性条件下，当我们对（实际）投资项目进行决策时，可以应用某种类型的分离原理。首先，经理选择投资项目来最大化公司价值，可以不管投资者的偏好。其次，为了最大

化预期效用，投资者可以通过借贷方式对“今天”和“明天”的消费进行配置。

* 在单期模型中，假设效用函数只依赖于投资组合的预期收益和收益的方差，模型所导出的结论是，资产份额依赖于预期超额收益、收益的协方差矩阵和投资者“风险规避”的程度。

* 只依赖于期末财富预期效用的单期模型一般无法得到闭型解。例外的情形是二次效用函数，或负指数效用函数加多维正态分布收益假设，这里的资产需求是预期超额收益的线性函数。对许多其他效用函数而言，资产需求的最优解必须用数值方法来进行计算。

附录：均值方差模型与效用函数

如果投资者追求投资组合期末财富的预期效用最大化，那么可以证明，在如下假设下，这等价于某个函数的最大化，函数是期末投资组合预期收益和方差的函数：

(a) 要么效用函数是二次效用函数；

(b) 要么收益服从正态分布（效用函数是凹函数）。

如果初始财富为 W_0，R_p 是随机的投资组合收益，那么期末财富和效用分别为

$$W = W_0\ (1 + R_p) \tag{A1}$$

$$U\ (W)\ = U\ [W_0\ (1 + R_p)] \tag{A2}$$

将 $U\ (R_p)$ 在 R_p 的均值（$=\mu_p$）附近进行 Taylor 级数展开后，我们就得到

$$U\ (R_p)\ = U\ (\mu_p)\ +\ (R_p - \mu_p)\ U'\ (\mu_p) + \frac{1}{2}\ (R_p - \mu_p)^2 U''\ (\mu_p)\ + \text{高阶项} \tag{A3}$$

因为 $E\ (R_p - \mu_p)\ = 0$，$E\ (R_p - \mu_p)^2 = \sigma_p^2$，所以对（A3）求数学期望就得到

$$E\ [U\ (R_p)]\ = U\ (\mu_p)\ + \frac{1}{2}\sigma_p^2 U''\ (\mu_p)\ + E\ (\text{高阶项}) \tag{A4}$$

如果效用函数是二次函数，那么除 U'' 之外的高阶矩都为零。如果收益服从正态分布，那么，当 n 为奇数时，$E\ [\ (R_p - \mu_p)^n]\ = 0$，当 n 为偶数时，$E\ [\ (R_p - \mu_p)^n]$ 只是方差 σ_p^2 的函数。因此，就情形（a）和（b）而言，$E\ [U\ (R_p)]$ 只是均值 μ_p 和方差 σ_p^2 的函数。这一结果在某些情形下是非常适用的。例如，可以证明，如果定义效用函数只是 μ_p 和 σ_p 的函数且为凹函数，那么（μ_p，σ_p）空间上的无差异曲线是凸的，正如正文中（图 4）所假设的那样。然而，除非我们假设效用函数是某个特殊的效用函数，否则无法知道 $E\ [U\ (R_p)]$ 与（μ_p，σ_p）之间的函数关系，从而不能确定资产的需求函数是否有解析的闭型解。

使用二次效用函数的一个问题是，在极乐点上，边际效用与财富水平负相关。资产收益服从正态分布的假设可能不是一种准确的表述，价格也可能变为负数。连续时间模型假设收益服从瞬时正态分布，这一假设使得问题变得相当容易处理（见 Cuthbertson 和 Nitzsche（2001a）），在衍生金融工具的定价中广泛使用这一假设。即使在高频数据下（如 tickdata），正态分布的实证有效性仍然处于争论之中，但结果的有用性取决于所讨论的问题（例如，对股票指数期权进行定价似乎就非常合理）。

第2章

金融中的统计基础

目的

* 考察对数正态分布和 Jensen 不等式。
* 讨论单位根、随机游走和协整之间的关系。
* 说明蒙特卡罗（Monte Carlo）模拟（MCS）和自助法（bootstrapping）的使用。
* 为了给出利率参数的最优预测值，指出如何将贝叶斯（Bayes）规则用于先念信息与数据估计的联合使用。

假设读者已经具备良好的本科层次的统计学和计量经济学基础知识，本章内容是所选择的几个主题的简略叙述，这些主题对金融经济学领域实证文献的理解至关重要。

2.1　对数正态分布与 Jensen 不等式

假设 Z 服从对数正态分布，$z=\ln Z\sim N(\mu_z,\ \sigma_z^2)$，其中 $\mu_z=E(\ln Z)$，$\sigma_z^2=E(\ln Z-E(\ln Z))^2$。从而 Z 的数学期望值为

$$EZ=Ee^z=e^{\mu_z+\frac{1}{2}\sigma_z^2} \tag{1}$$

方程（1）是 Jensen 不等式的一种特殊情形，将 Jensen 不等式应用于非线性函数的数学期望上就可以得到

$$E[f(Z)]\neq f(EZ) \tag{2}$$

这里的 $Z=\exp(z)$，但 $EZ\neq\exp(Ez)=\exp(\mu_z)$。为了说明这一点，假设 k 期远期利率是在 k 期之前对未来汇率的一个无偏预测：

$$E_tS_{t+k}=F_t^k \tag{3}$$

假设即期利率服从条件对数正态分布：

$$\ln S_{t+k}\equiv s_{t+k} \tag{4a}$$

$$s_{t+k}\mid\Omega_t\sim N[E_ts_{t+k},\ \mathrm{Var}_t(s_{t+k})] \tag{4b}$$

这里的 Ω_t 是 t 时刻的信息集合。由（1）就得到

$$\begin{aligned}E_tS_{t+k}&=E_t[\exp(s_{t+k})]\\&=\exp\left[E_tS_{t+k}+\frac{1}{2}\mathrm{Var}_t(s_{t+k})\right]\end{aligned} \tag{5a}$$

$$\ln(E_tS_{t+k})=E_ts_{t+k}+\frac{1}{2}\mathrm{Var}_t(s_{t+k}) \tag{5b}$$

在（3）中应用（5b），

$$f_t^k = E_t s_{t+k} + \frac{1}{2}\mathrm{Var}_t\ (s_{t+k}) \tag{6}$$

其中$f_t^k = \ln\ (F_t^k)$。因此，当即期利率服从条件正态分布时，远期利率的对数是未来即期利率（取对数）预期值的无偏预测。(6) 中增加的条件方差项经常被称为 Jensen 不等式项。在对远期利率的无偏性和后面章节中的其他问题进行讨论时，这一项就变得非常重要。

汇率可以用每单位外币的国内货币 S_t 或相反的方式来进行度量，因此

$$S_{t+1}^* = \frac{1}{S_{t+1}} \tag{7}$$

S_{t+1}^*用每单位国内货币的外币来进行度量。因为 (7) 是一个非线性变换，所以

$$E_t S_{t+1}^* = E_t\left(\frac{1}{S_{t+1}}\right) \neq \frac{1}{E_t\ (S_{t+1})} \tag{8}$$

从而 $E_t\ (1/S_{t+1}) \neq 1/E_t\ (S_{t+1})$ ——这是 Jensen 不等式的一种特殊情形，也是众所周知的 Siegel 矛盾（Siegel's paradox）。然而，对汇率的对数值 $s_{t+1} = \ln S_{t+1}$ 而言，下式是正确的

$$E_t\ (s_{t+1}) = -E_t\ (-s_{t+1}) \tag{9}$$

2.2 单位根、随机游走与协整

带飘移项（ $=\mu$）的随机游走是具有如下行为的单个随机序列 x_t

$$x_t = \mu + x_{t-1} + \varepsilon_t \quad \varepsilon_t \sim iid\ (0,\ \sigma_\varepsilon^2) \tag{10a}$$

这里的 *iid* 表示“独立同分布”。标准的计量经济检验是建立在变量平稳（也就是说，从宽泛意义上讲，总体的无条件均值为常数，方差和自相关系数只取决于序列滞后的时间长度）假设基础上的。(10a) 中序列 x_t 是非平稳的（即随机地向上或向下运动，几乎不会回到起点），我们称它是具有单位根的过程（即 $(1-\rho L)\ x_t = \varepsilon_t$，其中 $\rho = 1$）。单位根导致了随机趋势，因为根据 (10a) 有

$$x_{t+T} = T\mu + x_t + v_t \tag{10b}$$

$$E_t x_{t+T} = T\mu \tag{10c}$$

$$\mathrm{Var}_t\ (x_{t+T}) = T\sigma_\varepsilon^2 \tag{10d}$$

其中，$v_t = \varepsilon_{t+1} + \varepsilon_{t+2} + \cdots + \varepsilon_{t+T}$。即使 $\mu = 0$（即没有飘移项），序列从起点开始也会进行“随机游走”，当 $T \to \infty$ 时，序列的方差随着 T 的增加而增加。这就是随机趋势。x_t 的一阶差分 $\Delta x_t = \mu + \varepsilon_t$ 是平稳的（因为 ε_t 是平稳的），均值为常数。可以用扩展的 Dickey - Fuller 检验或 ADF 检验来对单个序列 x_t 的平稳性进行检验，ADF 检验是建立在 OLS 回归基础之上的

$$\Delta x_t = \mu + \gamma x_{t-1} + \sum_{i=1}^{m} \delta_i \Delta x_{t-i} + \varepsilon_t \tag{11}$$

ε_t 服从正态分布且 *iid*，通常记为 *niid* 或 $N\ (0,\ \sigma_\varepsilon^2)$。非平稳性的零假设为

$$H_0: \gamma = 0$$

平稳性的一种替代表示是 $H_0: \gamma < 0$（单尾检验）。$H_0: \gamma = 0$ 的 OLS “t 统计量”为

$$\hat{\tau}_\mu = \hat{\gamma}/se\ (\hat{\gamma}) \tag{12}$$

下标 μ 意味着数据生成过程（DGP）或“真实”模型为 (10a)。在零假设下，$\hat{\tau}_\mu$ 不服从学生 t 分布，因此它的 5% 的临界值必须用 MCS（见后面）来确定，该临界值依赖于 DGP 中 μ 的“真实”值和样本容量 n。例如，当 $u = 0$，$n = 50$（且 (11) 中滞后长度为 $m = 0$）时，5%（单尾）的临界值为 -2.92，而当 $\mu = 0.25$ 时，临界值为 -2.63（Patterson (2000)，表 6，p232）。当 $\mu \neq 0$ 时，$\hat{\tau}_\mu$ 的分布为渐近正态分布，5% 的临界值为 -1.65，但我们几乎不可能有足够多的数据来使用渐近临界值。当 DGP 中的 $\mu = 0$ 且 $n = 50$ 时，假定$\hat{\tau}_\mu$ 的值小于 -2.92，我们可以拒绝非平稳的零假设。上述检验方法有许多变体，但所有检验方法都遵循如上所述的普

遍原理（这个问题的优秀综述可以参考 Patterson（2000））。

变量平稳性的单位根零假设检验（问题）似乎有些“严重”。如果你不拒绝零假设，那么你就会接受如下结论：序列可能“游走”到正无穷或负无穷（即无条件方差与“时间”成比例），“冲击”具有持久效应。但是，如果你拒绝了零假设，那么序列协方差就是平稳的且不会“发散”。这种极端二分法的一种替代是使用分整（fractional integration）的思想。这里假设单变量序列遵循如下过程：$(1-L)^d x_t=\varepsilon_t$，其中 ε_t 是平稳的（但不必是白噪声）。如果 $d \geqslant 0.5$，那么这意味着序列是非平稳的，但没有必要具有发散行为。当 d 趋于 1 时，序列的非平稳性增加，一直类推下去，当 x_t 为单位根时，$d=1$。如果 $d \in [0.5, 1)$，那么 x_t 是方差非平稳的，但会出现均值反转，从而创新（innovation）对 x_t 没有持久效应。因此，我们可以通过对 d 的各种取值进行检验（Robinson（1994））来获得序列非平稳的程度及是否发散的一些想法。然而，这种方法在文献中并不流行。这是因为，I（1）变量之间的协整（而不是分整）意味着我们可以使用多变量方法来对经济理论进行检验（见下面）。

值得注意的是，单变量序列协整阶数的绝大多数检验（相对于“近似”单位根的替代检验方法而言）的检验功效非常的低，金融中许多感兴趣的时间序列都“近似”为单位根序列。因此，在对这一领域的许多“统计结果”进行解释时，人们必须使用他们自己的判断。

协整

进行差分后变成平稳序列（差分的次数最小为 d）的任何序列 x_t 都被称为“d 阶自积”序列，或简称为“I（d）”序列。随机游走为 I（1）序列。（注意，不是所有的非平稳序列都能通过差分的方法变成平稳序列。）

如果两个序列 x_t 和 y_t 都是 I（1）序列，（比方说）那么 x_t 和 y_t 的线性组合在一般情况下也为 I（1）序列：

$$y_t-\beta x_t=z_t \tag{13}$$

然而，可能出现的情形是，x_t 中的（随机）趋势被 y_t 中的趋势“抵消掉”，从而线性组合是平稳序列（具有常数 β）。如果这种“特殊情形”成立，那么称 x_t 和 y_t 是协整的两个序列。协整的一种简单检验方法是进行 OLS 回归

$$y_t=\hat{\alpha}+\hat{\beta}x_t+e_t \tag{14}$$

并用（11）对残差 e_t 进行平稳性检验（尽管由 MCS 得出的零假设 H_0: $\gamma=0$ 的临界值与上面所给出的临界值不同）。

Engle－Granger 表示定理

Engle 和 Granger（1987）证明了如下结论：如果两个（或多个）I（1）序列是协整的，那么动态误差修正模型（ECM）存在，线性组合 $(y-\beta x)_{t-1}$ 有助于预测 Δx_t 或 Δy_t，或对两者同时具有预测作用。为了用一个最简单的例子来说明这一点，假设（x_t，y_t）都是 I（1）序列，且具有协整关系，从而 $(y-\beta x)_t$ 是平稳序列。$(y-\beta x)_t$ 必然有一个平稳的自回归移动平均（ARMA）表示

$$(y_t-\beta x_t)=\rho_1(y-\beta x)_{t-1}+\rho_2(y-\beta x)_{t-2}+v_t \tag{15a}$$

$$\Delta x_t=\varepsilon_t \tag{15b}$$

我们在这里假设 $(y-\beta x)_t$ 服从 AR（2）过程，x_t 为随机游走（飘移项为零），v_t 和 ε_t 是白噪声过程。

由（15a）和（15b）可以直接导出误差修正模型 ECM

$$\Delta y_t=\delta_1\Delta y_{t-1}+\delta_2\Delta x_{t-1}+\lambda_y(y-\beta x)_{t-1}+\omega_t \tag{16}$$

其中，$\delta_1=-\rho_2$，$\delta_2=\rho_2\beta$，$\lambda_y=\rho_1+\rho_2-1$，$\omega_t=v_t+\beta\varepsilon_t$。因为 $(y-\beta x)_t$ 是平稳序列，所以 $\rho_1+\rho_2<1$，从而 $\lambda_y<0$。因此，如果 $y_{t-1}>\beta x_{t-1}$，那么下期的 $\Delta y_t<0$（其他条件相同），y 向其长期均衡 $y^*=\beta x$ 方向运动。

如果Δx_t为ARMA过程，那么Δx_t也可以表示与（15a）类似的误差修正模型，协整项相同都为$(y-\beta x)_{t-1}$，但系数λ_x不同。两个误差修正方程，一个是Δx_t的，另外一个是Δy_t的，合在一起就被称为双变量ECM（VECM）。接下来的实证研究问题是，λ_x和λ_y是同时非零，还是只有一个非零。如果y_t和x_t是协整的，那么λ_x和λ_y中至少有一个必须非零。Engle－Granger定理也意味着逆定理成立，即如果存在Δx_t或Δy_t（或两者）的误差修正表示，那么x_t和y_t是协整的。

构建两个（或多个）I（1）序列并应用Engle－Granger定理来推断预测性的方法在如下情形中被广泛使用：序列中的一个（或多个）为资产价格（如股票、债券或外汇价格）序列。这种方法至少在一个方向上可以进行预测，VECM模型可以用来考察单向或双向因果关系。值得注意的是，建立股票收益的这种预测性（通过协整和VECM模型）确实意味着，你在重复博弈中能获得正的利润，但它并不必然意味着这些利润足以弥补交易成本和补偿重复博弈中的潜在风险。本书余下部分将对“有效市场”中的这个问题进行详细讨论。

如果存在n（>2）个I（1）变量，那么至多存在（$n-1$）个协整向量，对$n>2$而言，这些协整向量不唯一。从原理上讲，每一个协整向量都可以出现在变量ECM模型中。Johansen（1988）的程序允许我们对VECM系统中协整向量的个数进行检验，也能对误差修正项的参数β的限制进行检验（现在，对不同的协整向量而言，β是一个向量）。

协整文献中许多检验统计量的小样本性质必须用蒙特卡罗模拟或自助法来获得（如Psaradakis（2001）），如果序列近似为单位根，那么这些统计量的功效性质通常非常糟糕（为了说明这一点，参见网上的程序）。

2.3 蒙特卡罗模拟与自助法

计量经济学中经常出现的情形是，我们知道估计量的性质及其统计分布的渐近分析结果，但不知道小样本下估计量有哪些性质。在这种情形中，蒙特卡罗（MCS）就变得非常有用。

Dickey－Fuller检验

为了着手进行分析，考虑$\hat{\tau}_\mu=\hat{\gamma}/se(\hat{\gamma})$的临界值，我们用MCS来获得$H_0: \gamma=0$的Dickey－Fuller“$t$统计量”值。下标$\mu$意味着数据生成过程（DGP）或“真实”模型为（10a）。假设我们从$\varepsilon_t \sim N(0, 1)$中抽出一个（比如说）$n=50$的随机样本。令（10a）中$\mu=0.25$，$x_0=0$，$\varepsilon_t$的50个随机值生成了零假设下的一个模拟数据序列。利用x_t的这50个模拟观察值进行Dickey－Fuller OLS回归$\Delta x_t=\hat{\mu}+\hat{\gamma}x_{t-1}\hat{v}_t$，保存$\hat{\tau}_\mu=\hat{\gamma}/se(\hat{\gamma})$的值。这是MCS的第一次运算。现在，我们根据$\varepsilon_t$的另一组随机抽取值（$t=1, 2, \cdots, 50$）来重复上面的操作，由此可以（在零假设$\gamma=0$下）得到一个大数目（比方说20 000）的$\hat{\tau}_\mu$值，将这些$\hat{\tau}_\mu$值画成柱状图。一个众所周知的结果是，$\hat{\gamma}$是向下偏斜的，因此$\hat{\tau}_\mu$并不是以0为中心（与零假设下的结论一样）。现在，我们来寻找$\hat{\tau}_\mu$柱状图中的切断点，该点使得拟合分布的左尾区域占整个区域的5%。正如上面所注意到的那样，这个值为$\hat{\tau}_\mu=-2.63$（对$\mu=0.25$，$n=50$而言），它就是单尾检验的临界值。接下来，针对各个样本长度、DGP中μ的各种可能取值和回归方程中各种滞后长度（见(11)），重复MCS来得到各种情景下的临界值（这些临界值也经常用一个简单的公式来进行表示：临界值取决于样本容量、μ的值等。这就是众所周知的反应面（response surface））。

收益回归

接下来，我们考察金融计量经济学中一个非常知名的问题，问题可以表示为：

$$y_t=\alpha+\beta x_{t-1}+u_t \tag{17a}$$

$$x_t = \theta + \rho x_{t-1} + v_t \tag{17b}$$

其中，$|\rho| < 1$，u_t 和 v_t 具有同时相关性，$\sigma_{uv} \neq 0$。然而，不同时点上的误差项服从独立同分布的正态分布（即每一个误差项都无异方差或序列相关性）。假设两个误差项服从二维正态分布，误差项的协方差矩阵为

$$\sum = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix} \tag{18}$$

y_t 可以取为第 t 期中资产价格的变化（即收益），x_{t-1} 是在 $t-1$ 期期末就已经知道的一些变量。如果 y_t 为股票的（单期）收益，那么 x_{t-1} 就可能是红利价格比（即红利收益）或账面市值比或收益率差。在债券市场上，y_t 可能是持有期收益或债券收益的变化，x_{t-1} 可能是收益率差。在外汇市场上，y_t 可能是（单期）即期外汇汇率的变化，x_{t-1} 是远期升水（forward premium）。在本书后面的部分中，在对有效市场假说进行检验的背景下，所有这些情形我们都将遇到。

我们在这里将分析限制为 y_t 是单期收益的情形。在考察多期收益时，我们会面临相同的问题。例如，如果使用的是重叠数据（如月度数据样本，而 y_t 是每年的价格变化），那么 u_t 序列相关，下面对标准估计量（如 GMM）的讨论也有类似结果。可以证明，（17a）中的 OLS 估计值 β 是一致的，但也是有偏的，偏差的根源是 u_t 和 v_t 之间存在相关性。背后的直觉如下。假设 $u_t > 0$，$\sigma_{uv} > 0$，从而 v_t 的平均值也大于零。v_t 的增加导致了 x_t 的上升，尽管（17a）中 u_t 从而 v_t 与 x_{t-1} 无关，u_t 还是与 x_t 相关，后者导致 OLS 估计值 $\hat{\beta}$ 是有偏的。

证明（17a）中 OLS 估计值 $\hat{\beta}$ 是最优线性无偏估计量（BLUE）的标准假设是

$$E\{u_t | \cdots, x_{t-2}, x_{t-1}, x_t, x_{t+1}, x_{t+2}, \cdots\} = 0 \tag{19}$$

因此，假设 u_t 与 x_t 所有的未来项和滞后项都无关。然而，在考虑 $E(u_t | x_t, x_{t+1}, x_{t+2}, \cdots)$ 时，如上所述，这个数学期望并不等于零，标准 OLS 结果不再成立。可以证明 $\hat{\beta}$ 是渐近一致的，且

$$t = \frac{\hat{\beta} - \beta}{se(\hat{\beta})} \sim N(0, 1) \tag{20}$$

因此，对 β 的假设检验在大样本情形中是"标准的"。注意，当 $n \to \infty$ 时，这个（伪）"t 统计量"的分布为 $N(0, 1)$。然而，在小样本中，OLS 估计值 $\hat{\beta}$ 是有偏的，$\hat{\sigma}_u^2$ 和 $se(\hat{\beta})$ 的 OLS 估计值也是有偏的，（20）中的"t 统计量"并不服从学生 t 分布（或正态分布）。当我们只有有限样本时，如果使用（20）的渐近结果，那么所做的推断是错误的。怎样才能对 OLS 估计量 $\hat{\beta}$ 的小样本分布进行考察呢？假设试图进行检验的零假设为 $\beta = 1$（这样选择是为了方便起见），y_t 是标准普尔 500 指数的月度（超额）收益，x_{t-1} 为上月末的红利收益（通常用 $t-2$ 到 $t-1$ 所付红利除以标准普尔 500 指数在 $t-1$ 期末的价格来对其进行度量）。

假设（17a）中 OLS 回归使用的是年度观察值的实际值数据，样本容量为 $n = 50$ 年，回归结果为 $\hat{\beta}_d = 2.2$，$se(\hat{\beta}_d) = 0.2$。对 $H_0: \beta = 1$ 使用（20）中的渐近统计量就得到

$$t_d = (2.2 - 1) / 0.2 = 6 \tag{21}$$

在 5% 显著性水平下的双尾检验中，明显要拒绝 H_0（因为临界值等于 ±1.96），$t_d = 6$ 的 p 值小于 0.001。为了考察零假设 $H_0: \beta = 1$ 下 OLS 统计量的小样本性质，我们进行 MCS，生成 x 和 y 的拟合数据。为此，我们需要参数 α，β，θ，ρ 和 $\sum$ 的值。

在进行第一次逼近时，我们能使用由"真实"数据集合对（17a）和（17b）同时进行回归所得到的这些参数的"OLS"估计值（即 SURE 估计值，这些估计值是一致的）。在有限样本中，尽管这些估计值是有偏的，但为 MCS 分析提供了一个合理的起点。（我们在后面会多次看到结论对这些参数变化的敏感性。）

假设我们所发现的 OLS 估计结果为 $\alpha = 0.2$，$\theta = 0.1$，$\rho = 0.8$，后者与观察到的价格红利

比的持续性（或匀滑性）相一致。根据（17a）和（17b）的残差，我们也得到了Σ中元素的估计值。

在整个 MCS 过程中，参数$\{\alpha, \theta, \rho, \Sigma\}$都是固定的，在选择“真实值”时$\beta=1$是固定的（通常由经济理论来给定）。MCS 包括如下步骤：

（ⅰ）从协方差矩阵为Σ的二维正态分布中抽取u_t和v_t的（n_1+n）个值；

（ⅱ）以$x_0=4\%$（数据中的平均红利价格比）为起点，迭代生成序列x_t；

（ⅲ）利用所生成的x序列和随机误差来生成y_t的（n_1+n）个值。

（ⅳ）扔掉x和y的前n_1个值，还剩下n个值；

（ⅴ）计算$\hat{\beta}^{(1)}$的 OLS 估计值和其他感兴趣的统计量，如渐近 OLS 标准差se（$\hat{\beta}^{(1)}$）；

（ⅵ）重复（ⅰ）～（ⅴ）10 000 次（$m=10\ 000$），得到$\hat{\beta}^{(i)}$和se（$\hat{\beta}^{(i)}$）（1，2，…，m）的 10 000 个值。

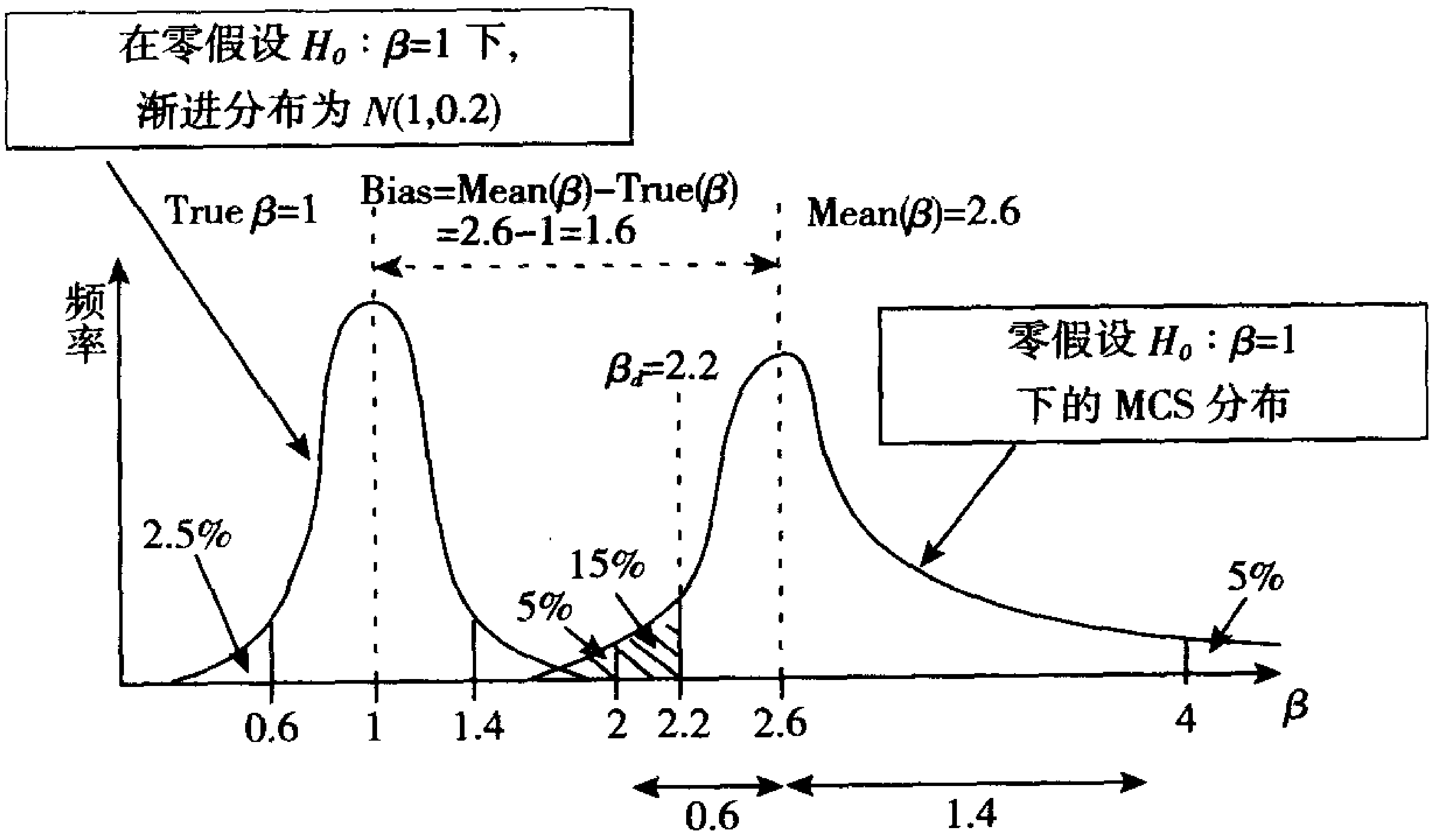


图 1　渐近分布和 MCS 分布

扔掉前n_1个数据的目的在于，使生成的数据与初始值x_0无关（一般而言，$n_1=150$就足够了）。固定的样本“长度”$n=50$是非常关键的，因为它代表的是现实世界中“小样本”的容量。重复“运算”MCS 的数目（$m=10\ 000$）不会影响估计量的性质，为了获得$\hat{\beta}^{(i)}$分布相当准确的估计结果，该数目必须足够地大（见图 1）。计算 10 000 个$\hat{\beta}^{(i)}$的均值

$$\text{mean}\ (\hat{\beta})\ =\ \sum_{i=1}^{m}\beta^{(i)}/m=2.6\ (\text{比方说}) \tag{22}$$

因此，当知道（17a）中$\beta=1$时，我们会发现，在样本容量为$n=50$的 MCS 中，β的 OLS 估计量的平均值为 2.6。（我们使该差异“这么大”的目的是为了说明：在对利率期限结构的预期假说进行检验时，这样大的差异完全可能发生，我们在后面也会看到这种现象，Bekaert、Hodrick 和 Marshall（1997））。β的 OLS 小样本偏差为：

$$\text{bias}\ (\hat{\beta}_{\text{OLS}})\ =\text{mean}\ (\hat{\beta})\ -\text{true}\ (\beta)\ =2.6-1=1.6 \tag{23}$$

使用实际数据所得到的“一个”（one shot）*OLS* 估计值为$\hat{\beta}_d=2.2$。但是，偏差解释了该值中的“1.6”，因此由实际数据所得到的β的一个更为准确的有限样本估计值为

$$\text{adj}\ (\hat{\beta}_d)\ =\hat{\beta}_d-\text{bias}\ (\hat{\beta}_{\text{OLS}})\ =2.2-1.6=0.6 \tag{24}$$

现在，adj（$\hat{\beta}_d$）＝0.6 的偏差调整值更接近于理论值 1，我们接受零假设的可能性也就更大（其他条件相同）。小样本的 adj（$\hat{\beta}_d$）与零假设$\beta=1$之间的“差距”现在仅仅是 0.4（绝

对值）：

$$\text{distance} = \text{adj}\ (\hat{\beta}_d)\ -\beta = 0.6 - 1 = -0.4 \tag{25}$$

现在对 $H_0: \beta = 1$ 进行如下“t 检验”：

$$\text{adj} - t = \frac{\text{adj}\ (\hat{\beta}_d)\ -\beta}{\text{se}\ (\hat{\beta}_d)} = \frac{0.6 - 1}{0.2} = -2 \tag{26}$$

在5%的显著性水平下我们恰好能拒绝零假设（双尾检验），因为由正态分布可知，（渐近）临界值为 -1.96。但这个结果仍具有误导性，因为我们在计算 se（$\hat{\beta}_d$）时使用的是有偏的 OLS 公式。我们也假设“adj - t”在小样本下的分布为 N（0，1），但这并不是真实的情形。能用 MCS 结果来对此进行改进吗？

当知道 $\beta = 1$ 时，（误差项的）方差 - 协方差矩阵 Σ 是真实数据结果的一个替代，图 1 给出了 MCS 10 000 个 $\hat{\beta}^{(i)}$ 值的分布，出于纯粹的教学目的，我们假设 $\hat{\beta}^{(i)}$ 的分布不是对称分布，因此均值是集中趋势的确切指标。在对结论进行敏感性检验时，将（23）和（24）中的 mean（$\hat{\beta}$）用中值来代替。但我们仍然将均值作为集中趋势的一个测度指标。

在图 1 中，我们会看到，$\hat{\beta}^{(i)}$ 的经验分布（在 mean（$\hat{\beta}$）附近）比渐近分布更为分散。事实上，根据 MCS，我们能计算出 $\hat{\beta}^{(i)}$ 的标准差：

$$\text{mean}\ (\hat{\beta}) = \sum_{i=1}^{m} \hat{\beta}^{(i)}/m \tag{27a}$$

$$\text{stdv}\ (\hat{\beta}) = \Big[\sum_{i=1}^{m} \{\hat{\beta}^{(i)} - \text{mean}\ (\hat{\beta})\}^2 / \ (m-1) \Big]^{1/2} = 0.4 \text{（比方说）} \tag{27b}$$

它是 OLS 估计值的渐近分布标准差 0.2 的 2 倍。

经常用这些报告的 MCS 标准差来替换解析的渐近值，因为在小样本下，我们认为前者对真实值的代表性要更强一些。注意，经验分布不是对称分布，标准差可能不是 β 估计值分散程度的一个有用度量指标。

如果 stdv（$\hat{\beta}$）＝0.4，那么在（26）中运用这个结果就得到 adj - t = 1，接受渐近正态分布假设意味着我们现在不能拒绝 $H_0: \beta = 1$。但是，一个明显的问题是，由 MCS 得到的 $\hat{\beta}^{(i)}$ 的有限样本分布并不是正态分布，因此在对 $H_0: \beta = 1$ 进行检验时，我们不能使用正态分布的临界值。我们所能做的是，将用实际数据获得的 β 估计值 $\beta_d = 2.2$ 与经验分布中的 p 值进行比较。

从图 1 中我们可以看到，当真实的 β 等于 1 时，mean（$\hat{\beta}$）＝2.6。因此，我们应用统计的方法来对由实际数据得到的 OLS 估计值 $\hat{\beta}_d = 2.2$ 与 mean（$\hat{\beta}$）＝2.6 之间的“多大差距”进行度量。（注意，2.6 - 2.2 = 0.4 与（25）中的“距离”相同。）

$\hat{\beta}_d = 2.2$ 的一个有用“统计度量”是 MCS 中 $\hat{\beta}^{(i)}$ 小于 2.2 的比例，它为 15%（图 1）。因此，蒙特卡罗分布（它包含了 OLS 的小样本偏差）意味着，给定 $\hat{\beta}_d = 2.2$，在零假设（$\beta = 1$）下，p 值为 15%，这表明我们不能拒绝 $\beta = 1$ 的零假设。

上述结果等价于将经验分布向左边移动（图 1），从而均值以 $\beta = 1$ 为中心，但标准差和分布的形状与 MCS 分布相同。基于实际数据的偏差调整估计值为 adj（$\hat{\beta}_d$）＝0.6，它的 p 值也为 15%（在将 MCS 分布向左移动之后）。

然而，用图 1 中的实际经验分布来进行计算要更为容易一些，我们在这里能计算出 $\beta_L = 2$ 和 $\beta_u = 4$ 的第 5 个百分位数和第 95 个百分位数。注意，这些值在 mean（$\hat{\beta}$）＝2.6 附近是不对称的，因为我们已经假设蒙特卡罗分布不是对称分布。在 90% 的置信水平下，我们可以说，如果 true（β）＝1，那么在有限样本中，由实际数据集合（n 个观察值）得到的 β 在 2 到 4 之间的 OLS 估计值并不能拒绝零假设 $\beta = 1$。因此，如果 true（β）＝1，那么实际数据的 OLS 估

计值 $\beta_d \leqslant 1.99$ 被观察到的概率小于 5%。换一种叙述方式是，假设 OLS 的估计值为 1.99，对 $H_0: \beta = 1$ 而言，经验 p 值为 0.05（尾部更低）。

尽管上述分析都是根据 β 的 MC 分布来进行的，我们还是可以用相同的方式来画出由 10 000次 MCS 所得到的“t 统计量” $\tau^{(i)} = (\hat{\beta}^{(i)} - 1) / \text{se}(\hat{\beta}^{(i)})$ 的 MC 分布（其中估计量 $\hat{\beta}$ 和 se（$\hat{\beta}$）为一致估计量）。因此，我们可以得到 $\tau^{(i)}$MC 分布的上下 2.5% 的尾部临界值，可能是 +2和 +6.2（比方说）。使用真实数据所得到的 t 统计量值为 $t_d = (2.2 - 1) / 0.2 = 6$，因此在“$t$ 统计量” MC 分布的基础上，在有限样本下，我们不能拒绝零假设 $\beta = 1$。相对于 $\hat{\beta}$ 的分布本身而言，人们更偏好使用 MCS（和自助法，见后面）所获得的“t 统计量”分布，因为后者具有更好的统计性质。称 t 统计量为“枢轴”统计量，这意味着其分布不依赖于诸如误差项方差之类的“冗余参数”。

解析表达式

在某些情形中，求得某个特殊模型的某个特殊估计值（如 OLS）的偏差、标准差和 p 值（零假设下）的解析表达式是可能的（需要付出许多努力）。在（17a）和（17b）的情形中，OLS 偏差为

$$E\hat{\beta} - \beta = \frac{\sigma_{uv}}{\sigma_v^2}\left[E\hat{\rho} - \rho\right] \tag{28}$$

其中，近似地有

$$E\hat{\rho} - \rho \approx -(1 + 3\rho)/n \tag{29}$$

因此，用 ρ_{uv} 表示 u 和 v 之间的相关系数，运用 $\sigma_{uv} = \rho_{uv}\sigma_u\sigma_v$ 就得到

$$E\hat{\beta} - \beta = -\frac{\rho_{uv}\sigma_u}{\sigma_v}\left(\frac{1 + 3\rho}{n}\right) \tag{30}$$

$\rho_{uv} < 0$（>0）表示 β 的 OLS 估计值具有向上（下）的偏差。如果 $\rho_{uv} = 0$，那么偏差为零（标准 OLS 的情形），大样本情形下（$n \to \infty$）偏差会消失。偏差越大，ρ（即 x_t 的持久性或自相关性）越大，相对于 σ_v 的 σ_u 越大。

自助法

在 MCS 中，假设已经知道误差项的联合分布（如多维正态分布）。然而，使用真实数据所得到的残差分布与正态分布之间有很大的差距，值得注意的是，残差可能是真实误差的有偏估计，因为 OLS 估计值 $\hat{\beta}$ 和 $\hat{\rho}$ 在小样本中是有偏的。但使用真实数据的残差能提供引导方面的所有信息。假设残差独立同分布，但不服从正态分布。因此，根据残差集合（替换 MCS 中所生成的多维正态误差），我们能重复使用随机抽取（要进行替换）MCS 的整个过程。这种方法称为自助法。自助法每进行一次运算后就得到 $\hat{\beta}^{(i)}$ 和 $\tau^{(i)}$ 的 OLS 估计值及其经验（自助）分布。例如，如果残差具有肥尾（与正态分布相比），那么我们可能会发现，用自助法获得的 $\hat{\beta}^{(i)}$ 和 $\tau^{(i)}$ 的经验分布尾部比用多维正态分布下 MCS 方法获得的分布尾部要厚。$\tau^{(i)}$ 的自助法分布也经常用于推断，因为它比 $\hat{\beta}^{(i)}$ 的分布“行为更好”（即是“枢轴”的）。

在自助法中，我们在（17a）和（17b）中配对地抽取 n 个残差（$\hat{u}_t$，$\hat{v}_t$）。如果 Cov（$\hat{u}_t$，$\hat{v}_t$）$\neq 0$，那么这种操作保留了残差中的任何同时相关性。为 x_t 和 y_t 拟合的或“自助的”时间序列在自助的每一次运算过程中都不同，主要有两个方面的原因。第一，抽取 n 个 $\hat{v}_t$ 值的顺序确定了 x_t（使用递归方法），自助法每一次计算的抽取顺序都不同。第二，因为我们“用替换的方法进行抽取”，所以实际数据的特殊残差对，比方说（$\hat{u}_{10}$，$\hat{v}_{10}$），在自助法的抽取过程中被抽取的次数可能不止一次。

如果残差 $\hat{u}_t$ 或 $\hat{v}_t$ 不是独立同分布，那么就必须对自助法进行修正。例如，如果残差 $\hat{u}_t$ 服

从 MA（1）过程，那么为了保留方程间的同时相关性，$\hat{u}_t$ 的每一次“自助抽取”都涉及相临的对（$\hat{u}_t$，$\hat{u}_{t-1}$）和（$\hat{v}_t$，$\hat{v}_{t-1}$）的相应值（Politis 和 Romano（1994））。如果 u_t 是 AR（1），那么为了充分地表示这种自相关模式，残差要从一个“块”中随机选择（这并不奇怪，这种方法被称为分块自助法——Li 和 Maddala（1997））。一种替代方法是，运用 Cochrane - Orrcutt 变换对（17b）进行重新估计，由此得到的“新”残差不具有自相关性，可以按照通常的方式用于自助（这就是“筛选自助法”（sieve bootstrap）背后的基本思想——Buhlmann（1997））。一般而言，通过一些技巧，我们能设计一些方法来生成残差的重复样本，从而使所生成的残差与从实际数据中发现的残差更好地“匹配”。

MCS 和自助法明显的局限在于，对特殊的模型、零假设和样本容量而言，这些结论才成立。然而，敏感性分析给出了各种情景下的一些结果（如假设过程 x_t 是一个单位根过程），“反应面”给出了各个样本容量（和 ADF 检验中诸如 μ 之类的“冗余参数”的各种可能取值）的临界值。对经济模型进行检验时，检验功效的增加导致这些技巧被更加广泛地使用。

MCS 的另外一种用法（自助法就用得更少一些）是用于建立理论模型、使用合理的参数估计值对模型进行校准、从模型的误差项（假设服从一个已知的分布（如多维正态分布））中随机地抽取样本来生成拟合数据。因此，生成的拟合数据与已知理论模型是一致的。根据理论模型重复生成拟合数据，由此可以对一些关键变量（如资产收益、消费等）的统计性质进行考察。统计性质可能包括均值、标准差、偏度、峰度、自相关性、交叉相关性及“拟合”数据序列间的回归结果。如果拟合数据的“矩”在很宽的范围内与真实数据的“矩”匹配，那么这说明理论模型与“典型事实”是一致的。使理论模型有效的方法经常被称为模型校准（calibration），在后面的章节中我们将遇到很多这样的例子。

2.4　贝叶斯学习

投资者有很多可供选择的方法来了解某个特殊随机过程的参数（如收益 R_{t+1} 是价格红利比当前值 $(D/P)_t$ 的函数，从而 $R_{t+1}=\alpha+\beta(D/P)_t+\varepsilon_{t+1}$）。在理性预期假设下，我们假设投资者知道参数 $\theta=(\alpha,\beta)$ 的准确值，并以确定的 θ 值来对 R_{t+1} 进行预测。很明显，这是一个很强的假设。一种替代方法是假设投资者根据新到的数据更新自己的看法，因此他们承认当前与 θ 有关的知识具有不确定性。贝叶斯学派假设投资者对参数 θ 拥有先念估计（如根据经济理论），随着样本数据的增加，他们对先念估计值进行更新，从而获得 θ 的后念估计值。我们在后面的章节中要遇到的许多模型都涉及贝叶斯学习，因此，我们来勾画一下这种方法的主要特征。

最简单的情形是考虑一个单参数分布，我们假设这个单参数分布为泊松分布。在注意到伦敦地铁新的组织变化后，假设你现在正呆在家里，你认为高峰时期伦敦地铁的平均到达时间要么是 $\lambda=2$ 要么是 $\lambda=4$（分钟）。假设这些信念成立的先念概率分别为：

$$P(\lambda=2)=0.8 \quad 或 \quad P(\lambda=4)=0.2$$

因此，在你先念观念给定的条件下，$\lambda=2$ 正确的可能性是 $\lambda=4$ 的 4 倍。现在，你奔向离你家最近的地铁站，列车在 $X=6$ 分钟后到达。在你观察到 $X=6$ 之后，直觉告诉你，$\lambda=4$ 发生可能性比 $\lambda=2$ 要高。也就是说，你的后念概率不同于先念概率。泊松分布的概率密度函数（p. d. f.）为

$$f(\lambda,x)=\frac{e^{-\lambda}\lambda^{x}}{x!} \tag{31}$$

因此

$$P(X=6\mid\lambda=2)=0.012 \tag{32a}$$

$$P(X=6\mid\lambda=4)=0.104 \tag{32b}$$

给定 $X=6$，$\lambda=2$ 的后念概率为

$$\begin{aligned} P(\lambda=2 \mid X=6) &= \frac{P(\lambda=2,\ X=6)}{P(X=6)} \\ &= \frac{P(\lambda=2)\ P(X=6 \mid \lambda=2)}{P(\lambda=2)\ P(X=6 \mid \lambda=2) + P(\lambda=4)\ P(X=6 \mid \lambda=4)} \\ &= \frac{0.8\ (0.012)}{0.8\ (0.012) + 0.2\ (0.104)} \\ &= 0.316 \end{aligned} \tag{33}$$

类似地，

$$P(\lambda=4 \mid X=6) = 0.684$$

因此，在观察到 $X=6$ 之后，先念概率 $P(\lambda=2)=0.8$ 减少到 $P(\lambda=2 \mid X=6)=0.316$ 的后念概率，而先念概率 $P(\lambda=4)=0.2$ 却增加到 $P(\lambda=4 \mid X=6)=0.684$ 的后念概率。

如果 $\theta=\lambda$ 的可能取值超过两个，贝叶斯学派必须对参数 θ 所有可能的先念值指定一个先念概率密度函数 $h(\theta)$。此外，一般情况下，我们有多个观察值来计算后念概率。假设 Y 是相对于参数 θ 而言的一个统计量，概率密度函数为 $g(y \mid \theta)$，其中 Y 可能是样本数据 $X=(X_1,\ X_2,\ \cdots,\ X_n)$ 的样本均值 $\bar{X}$，$g(y \mid \theta)$ 可能是正态分布（见下面）。统计量和参数的联合概率密度函数 $k(y,\ \theta)$ 为

$$k(y,\ \theta) = g(y \mid \theta)\ h(\theta) \tag{34}$$

Y 的边际概率密度函数为

$$k_m(y) = \int_{-\infty}^{\infty} h(\theta)\ g(y \mid \theta)\ \mathrm{d}\theta \tag{35}$$

给定 $Y=y$，参数 θ 的条件概率密度函数就是 θ 的后念概率密度函数，记为 $k(\theta \mid y)$：

$$k(\theta \mid y) = \frac{k(y,\ \theta)}{k_m(y)} = \frac{g(y \mid \theta)\ h(\theta)}{k_m(y)} \tag{36}$$

θ 的后念概率密度函数只取决于观察到的数据 Y。注意，如果 $h(\theta)$ 为一个常数，从而后念概率密度函数为均匀分布，那么这被称为是无信息先念。贝叶斯学派将 θ 视为一个随机变量，θ 的最优预测（猜测）取决于损失函数。例如，如果 Z 为一随机变量，“b”是对 Z 的猜测，损失函数为 $E[(Z-b)]^2$，那么最小化损失函数后得 $b=EZ$。因此，如果对 θ 进行猜测的 $\omega(y)$ 的损失函数为 $E[\theta-\omega(y)]^2$，那么 θ 的最优（贝叶斯）预测就是条件均值

$$\omega(y) = \int_{-\infty}^{\infty} \theta k(\theta \mid y)\ \mathrm{d}\theta \tag{37}$$

类似地，如果损失函数是误差的绝对值 $\mid \theta-\omega(y) \mid$，那么后念分布的均值就是 θ 的贝叶斯估计值。

或许一个简单的例子可以为贝叶斯定理的解释提供更多的直觉。假设 $Y=\bar{x}$ 是容量为 n 的随机样本的均值，因此 $x=\{x_1,\ x_2,\ \cdots,\ x_n\}$，其中

$$x_i \sim N(\theta,\ \sigma^2)$$

假设知道 σ^2 的确定性值。（因此，我们不必使用贝叶斯方法来更新 σ^2 的值。）标准的统计结果是，$g(y \equiv \bar{x} \mid \theta)$ 是如下形式的概率密度函数为

$$g(y \equiv \bar{x} \mid \theta) = N(\theta,\ \sigma^2 n) \tag{38}$$

假设我们对 θ 进行先念猜测，并运用先念概率密度函数 $h(\theta)$，假设该函数为正态分布的密度函数

$$h(\theta) = N(\theta_0,\ \sigma_0^2) \tag{39}$$

因此，后念概率密度函数 $k(\theta \mid y)$ 为

$$k(\theta \mid y) \propto \frac{1}{\sqrt{2\pi}\ (\sigma/\sqrt{n})} \frac{1}{\sqrt{2\pi}\ (\sigma_0)} \exp\left[\frac{-(y-\theta)^2}{2\ (\sigma^2/n)} - \frac{(\theta-\theta_0)^2}{2\sigma_0^2}\right] \tag{40}$$

现在，去掉与 θ 无关的常数项（即包含只与 y 有关的那些项）：

$$k(\theta \mid y) \propto \exp\left[-\frac{(\sigma_0^2+\sigma/n)\ \theta^2-2\ (y\sigma_0^2+\theta_0\sigma^2/n)\ \theta}{2\ (\sigma^2/n)\ \sigma_0^2}\right] \tag{41}$$

进行配方（并去掉与 θ 无关的那些项）后，上式变为

$$k(\theta \mid y) \propto \exp\left[-\frac{(\theta-m)^2}{\mathrm{Var}}\right] \tag{42}$$

其中，

$$\begin{aligned} m &= \frac{y\sigma_0^2+\theta_0\ (\sigma^2/n)}{(\sigma_0^2+\sigma^2/n)} \\ &= \left[\frac{\sigma_0^2}{\sigma_0^2+\ (\sigma^2/n)}\right]y+\left[\frac{\sigma^2/n}{\sigma_0^2+\ (\sigma^2/n)}\right]\theta_0 \end{aligned} \tag{43a}$$

$$\mathrm{Var}=\frac{(\sigma^2/n)\ \sigma_0^2}{\sigma_0^2+\ (\sigma^2/n)} \tag{43b}$$

因此，后念概率密度函数为正态分布的密度函数，均值为 m，方差为 Var。如果假设损失函数是误差的平方，那么 θ 的贝叶斯估计值就是后念分布的均值，由（43a）得出。后念均值是样本均值 $y=\bar{x}$ 与先念均值 θ_0 的加权平均，权重取决于先念方差 σ_0^2 和样本方差 σ^2/n（σ^2 已知）的相对大小。因此，先念估计值 θ_0 朝样本估计值 $\bar{x}$ 的方向“收缩”，随着 n 的增加，后念均值更靠近于 $\bar{x}$。如果上述损失函数为 $|\omega(y)-\theta|$，那么贝叶斯估计应是后念分布 $k(\theta \mid y)$ 的中值。

在上面的例子中，我们假设已经知道条件分布 $g(y\equiv\bar{x} \mid \theta)$，因为这是一个标准的统计结果（即样本均值是方差为 σ^2/n 的总体均值的极大似然估计值）。一般而言，我们没有必要知道这一结果，因此，在给定样本观察值 $\{x_1, x_2, \cdots, x_n\}$ 的条件下，我们用如下似然函数来替换 $g(y \mid \theta)$

$$L(\theta)=f(x_1 \mid \theta)\ f(x_2 \mid \theta)\ \cdots f(x_n \mid \theta)$$

在给定 θ 的情形下，它是 $\{x_1, x_2, \cdots, x_n\}$ 的联合概率密度函数。因此，后念概率密度函数为：

$$k(\theta \mid x_1, x_2, \cdots, x_n) \propto h(\theta)\ L(\theta) \tag{44}$$

所以，θ 的贝叶斯估计值，比方说为 $\omega(x_1, x_2, \cdots, x_n)$，是上述后念分布的一些特征，如均值或中值。（基础统计和贝叶斯定理的一个精彩入门读物参见 Hogg 和 Tanis（1993），本节的很大一部分就来自于该书。）

注意，贝叶斯定理并不是模拟学习行为的唯一方式。例如，如果愿意假设代理人对参数没有任何先念信息，那么他可能会使用一些相对简单的更新程序如递归 OLS。如果总体真实的参数值为常数，那么随着所获数据的增加，递归 OLS 的估计值最终会向真实参数值收敛。如果真实参数值本身具有时变性，且我们知道该时变性的随机性质（如参数服从随机游走），那么随着所获数据的增加，对参数进行更复杂的最优更新通常是可能的（如使用卡尔曼滤波（Kalman filter））。处理连续时变参数的一种替代方法是假设真实参数在任意状态下（如高通货膨胀率区间或低通货膨胀率区间）为常数，但是，不同状态所对应的常数不同。因此，投资者面临一个过滤问题，因为在他进行预测之前，他必须先计算出参数落入某个特殊状态的概率。对区制转换马尔可夫（Markov）过程而言，可以用 Hamilton（1994）的滤波方法来解决这个问题。上面的所有方法都是用来模拟学习行为的，正如我们在后面章节中即将看到的那样。

2.5 小结

* Jensen 不等式意味着随机变量非线性函数的数学期望不等于数学期望的函数，$E[f(Z)] \neq f(EZ)$。对某些分布而言，如对数正态分布 $\ln Z \sim N(\mu(\ln Z), \sigma^2(\ln Z))$，我们能得到期望值的显式表达式：$EZ = \exp(\mu + (1/2)\sigma^2)$。

* 当数据序列存在单位根时，一些统计检验性质是“非标准的”。如果两个或多个序列存在单位根，那么它们的线性组合可能是平稳序列——称这些序列是协整的。协整序列是平稳的，因此“标准”统计结果适用。如果两个（或多个）序列是协整序列，那么误差修正项（即滞后的协整向量）至少是某一个变量的“Granger 原因”。

* MCS 和自助法可用来确定有限样本下检验统计量的临界值，我们无法获得这些统计量的解析解。为了验证模型，理论模型的随机模拟生成了性质与真实数据相似的拟合数据。这就是模型校准。

* 贝叶斯定理可以用来将参数的先念信息与样本估计值最优地结合起来，这就得出了参数只依赖于观察数据和先念猜测的后念分布。随着数据的增加，后念分布的均值相对而言更多地由样本数据来决定，先念猜测的权重越来越小。因此，当投资者根据新获得的数据来对未知参数值的观念进行更新时，他们用贝叶斯定理来对“理性”学习过程进行建模。

第3章

有效市场假说

目的

- 证明 EMH（有效市场假说）意味着交易风险资产（平均而言）无法获得超额利润。
- 说明 EMH 对投资分析师、兼并收购、资本充足率和资本成本的意义。
- 在 EMH 背景下，引入公平赌博、鞅和随机游走的概念。
- 概述 EMH 的各种实证检验方法。

在英国，人们习惯地将电视节目称为“价格就是正确的”（The Price is Right），竞争者不得不试着猜测耐用消费品正确的市场价格。EMH（efficient markets hypothesis）的速记就是“价格就是正确的”的很好体现，这里的价格指风险资产的价格。资产的公平价值代表的是资产未来收益的 DPV（折现值），在“有效市场”上，市场价格始终等于“公平价值”。

一般而言，当经济学家谈到资本市场有效的时候，他们通常将资产价格和收益看作是由竞争市场上理性交易者的供给和需求结果来决定的。这些理性交易者迅速地吸收那些决定资产价格和收益的相关信息（如未来红利前景），并相应地调整价格。因此，在获得信息方面，个体并没有不同的比较优势。从而，在这样的世界里，没有任何机会来获得比股票风险的公平报酬还要高的收益。简而言之，交易的超额利润应等于零。因此，代理人有效地加工信息，直接将这些信息嵌入到股票价格中去。如果当前和过去的信息被直接嵌入到当前价格中，那么只有新信息或“消息”（news）才能导致价格变化。因为根据定义，消息是无法预测的，所以价格变化（或收益）是无法预测的：t 时刻及之前的任何信息都无助于改进收益的预测（或者，等价地说，个体无法减小预测误差）。预测误差与先前信息独立的性质就是众所周知的正交性质（orthogonality property），在对有效市场假说进行检验的过程中，它是被广泛使用的一个概念。

3.1　概述

在 EMH 下，股票价格 P_t 已经包含了所有的相关信息，t 时刻与 $t+1$ 时刻之间价格变化的唯一原因是“消息”或未预料到事件的到来。因此平均来看，预测误差（也就是说 $\varepsilon_{t+1} = P_{t+1} - E_t P_{t+1}$）应等于零，且与任何信息集合 Ω_t 不相关，这里的 Ω_t 是进行预测时所能获得的信息。经常称后者为 EMH 的理性预期（RE）部分，可以将其表示为

$$P_{t+1} = E_t P_{t+1} + \varepsilon_{t+1} \tag{1a}$$

$$E_t\ (P_{t+1} - E_t P_{t+1})\ = E_t \varepsilon_{t+1} = 0 \tag{1b}$$

$E_t \varepsilon_{t+1} = 0$ 的含义是，P_{t+1} 的预测是无偏的（也就是说，平均来看，实际价格等于预期价格）。注意，ε_{t+1} 也可以（粗略地）描述为 t 与 $t+1$ 之间持有股票预料不到的利润（或损失）。在

EMH 下，平均来看，预料不到的利润必须等于零，这可以用（1b）来表示。

“预测误差必须与 t 时刻（及之前）所获得的信息 Ω_t 相独立”的观点就是众所周知的正交性。可以证明，如果 ε_t 序列相关，那么这就违背了正交性。误差项序列相关的一个例子是一阶自回归 AR（1）：

$$\varepsilon_{t+1}=\rho\varepsilon_t+v_{t+1} \tag{2}$$

其中，v_{t+1}是（白噪声）随机变量（根据假设，它与 t 时刻的信息 Ω_t 相独立）。预测误差 $\varepsilon_t=P_t-E_{t-1}P_t$ 在 t 时刻是已知的，因此它是 Ω_t 的一个部分。方程（2）意味着，本期预测误差对下期预测误差 ε_{t+1}具有预测效应，但是，根据（1），后者有助于未来价格的预测。这就违背了 EMH。

我们能更直接地看到，为什么 ε_t 的序列相关性意味着 t 时刻的信息有助于预测 P_{t+1}，如下所述。将方程（1a）滞后一期并乘以 ρ：

$$\rho P_t=\rho\ (E_{t-1}P_t)\ +\rho\varepsilon_t \tag{3}$$

（1a）减去（3）并运用 $v_{t+1}=\varepsilon_{t+1}-\rho\varepsilon_t$ 就得到

$$P_{t+1}=\rho P_t+\ (E_tP_{t+1}-\rho E_{t-1}P_t)\ +v_{t+1} \tag{4}$$

由（4）可以看出，当 ε 序列相关时，明天的价格依赖于今天的价格，于是，根据今天的信息，明天的价格是（部分）可预测的。（注意，括号中的项是预期的变化，它是不可预测的。）因此，ε“无序列相关”的假设包含在 EMH 假设之中，EMH 假设今天获得的信息对明天的股票价格预测没有丝毫用处（即正交性）。

注意，EMH/RE 假设对 ε_t 分布二阶矩和高阶矩的形式施加了一些限制。例如，在不违背 RE 的前提下，ε_{t+1}的方差（记为 σ_{t+1}^2）与其过去值 σ_t^2 相关。（这就是 ARCH 过程。）RE 只对 ε_t 的一阶矩（即期望值）施加了限制。

有效市场假说经常用于股票收益 R_t 的讨论，它意味着人们通过买卖股票不可能获得超额利润。因此，与（1）相似的方程可以应用到股票收益上去。实际收益 R_{t+1}有时会高于预期收益，有时会低于预期收益，但是，平均来看，预料不到的收益或预测误差 ε_{t+1}为零：

$$\varepsilon_{t+1}=R_{t+1}-E_tR_{t+1} \tag{5}$$

其中，$E_t\varepsilon_{t+1}=0$。为了对 EMH 进行检验，我们需要投资者如何确定预期（或必要）收益的模型，这个模型应建立在理性行为（不知如何定义）基础之上。就当前而言，我们假设有一个非常简单的模型：

（ⅰ）股票不支付红利，从而预期收益就是由价格变化所引起的预期资本收益；

（ⅱ）只要预期（必要）收益为常数，投资者就愿意持有股票。

因此

$$R_{t+1}=k+\varepsilon_{t+1} \tag{6}$$

ε_{t+1}为白噪声，且与信息集合 Ω_t 独立。考虑风险资产的必要收益 k，必要收益包括无风险利率 r 和风险溢价 rp 两个部分（即 $k=r+rp$），在方程（6）中我们假设这两个部分在时间轴上都是常数。因为股票不支付红利，$R_{t+1}=\ (P_{t+1}-P_t)\ /P_t\approx\ln\ (P_{t+1}/P_t)$，所以方程（6）意味着

$$\ln P_{t+1}=k+\ln P_t+\varepsilon_{t+1} \tag{7}$$

方程（7）中价格 P 取对数后是带飘移项 k 的随机游走。注意，如果无风险利率 r 和风险溢价 rp 都是常数且红利为零，那么股票价格（取对数）在 EMH 下遵循随机游走。在实证研究中，经常要对 $t+1$ 时的“价格”进行调整，使其包含 t 到 $t+1$ 之间所支付的红利。“股票价格服从随机游走”的说法通常适用于“包含红利的价格”。在有些实证研究中，研究人员认为，股票收益主要取决于资本利得（或损失），因此他们使用不包含红利的价格。

就相对平静时期（排除像 1987 年 10 月和 2000—2003 年那样的“股灾时期”）上股票价格的每日变化而言，假设风险溢价为常数似乎是合理的。然而，当我们对股价的每日变化进行考察时，通常会发现，误差项是序列相关的，一周之内不同交易日的收益也不同。特别地，星期一的价格低于星期五的价格，这就是众所周知的“周末效应”。我们也会发现，对一些股票而

言，一月份的每日价格变化与其他月份不同。“周末效应”和“一月效应”都是明显可以预测的事件。因此，从预测的角度来看，股票收益取决于所获得的信息（如所考察的那一天是一周内的哪一天）。因为收益部分可预测，所以在风险溢价为常数的假设下，这是对 EMH 的一种违背。然而，在“现实世界”中不可能出现的情形是，这种可预测性意味着投资者能获得超额利润，因为我们需要考虑交易成本。

由上述讨论明显可以看出，为了对 EMH 进行检验，我们需要一个如何决定均衡（必要）收益的经济模型。我们对代表性投资者是否有效地使用信息进行检验的前提条件是，所选择的模型能正确解释预期收益。拒绝有效市场假说要么是因为我们使用了错误的均衡“定价模型”，要么是因为代表性投资者的真实情况是没能有效地使用信息。

正如上面所注意到的那样，描述 EMH 的另外一种方式是，在有效市场上，投资者不可能获得超额利润。在 EMH 下，投资者在每种证券上所获得的收益恰好能弥补证券的风险和交易成本。因此，在 EMH 下，交易股票没有任何获取超额利润的机会。经常称后者为“公平赌博”性质。

3.2 EMH 的含义

股票收益由竞争市场上理性投资者的行为来决定，均衡收益是所有公开信息的反映，这种观点在金融经济学家中获得广泛支持。股票价格也是基础价值（即未来红利的 DPV）的反映，这一更强的观点也获得了广泛支持。那么 EMH 对股票市场有什么样的含义呢?

就风险规避型投资者而言，EMH 意味着，他可以采用“买入并持有”（buy and hold）策略。他可以通过持有市场投资组合（或模拟市场投资组合的30种左右的股票）的方式来分散风险。Andrew Carnegie 的建议“将所有鸡蛋放在一个篮子里面，并对篮子进行观察”是应该避免的。如果 EMH 正确，那么投资分析师的作用将非常有限。举个例子来说，他们的作用包括：

（a）对模拟市场投资组合的30种左右股票的选择提供建议。

（b）改变财富中每种资产的持有比例，这个持有比例反映的是资产在投资组合中的市场份额，它会随着时间演变而发生改变。

（c）根据税收变化来改变投资组合（例如，如果红利税比资本所得税高很多，那么对高所得税的支付者而言，从边际角度来看，向低红利、高预期资本收益资产方向调整可能是最优的）。

（d）为了使买卖的交易成本达到最低，要“货比三家”。

在 EMH 下，投资分析师使用公开信息是不能成为“选优胜者”（pick winner）的，因此“主动型”投资管理人员是多余的。我们甚至可以走得更远一些：个体投资者的操作就是简单地购买“被动型”指数基金（如共同基金或信托公司），这些基金对诸如标准普尔500等特殊类型的市场指数进行跟踪，交易成本也很低（如每年低于1%）。诸如投资管理人员之类的实践家是不会爽快地接受这种观点的：在竞争性很高的有效市场上，他们的技能在很大程度上是一种多余。然而，他们却经常支持市场“有效”的观点。但他们所使用的术语“有效”通常是指，股票市场不受政府的干预（如，零印花税、对交易头寸和资本充足率实行最少的管制）。一个自相矛盾的结果是，主动型管理人员确实有助于价格迅速吸收信息，因此，即使他们没有获得（经风险调整后的）超额收益，他们的交易活动也确实有助于市场有效（Grossman 和 Stiglitz（1980））。

值得注意的是，绝大多数个体和机构不会持有包含所有可交易资产的“市场投资组合”。这就要求除美国居民以外的绝大多数投资者应主要持有外国证券（即绝大多数公司为外国人所有）。此外，绝大多数共同基金或信托公司进行过专门研究，它们针对特殊部门（如银行）

或特殊地理区域（如日本的股票）来发行基金。这种操作存在营销方面的原因。如果投资银行同时运营多只基金，那么至少存在一只基金，他们吹嘘这只基金能“击败市场”（或其他竞争对手）。

接管、合并与金融机构

现在转到一些公共政策问题上去。假设股票市场为实际资源（如固定投资）的配置提供了“正确的”信号。只有很小一个部分的公司投资是由新股发行来进行融资的（如在英国，总的来看为4%），因而股票市场上上市公司的平均收益率为风险调整后的“权益融资成本”提供了一个合理的测度。后者可以用来对所有股份公司实际投资项目（如投资评估）的未来预期利润进行折现。然而，如果股票价格不能反映基本层面的情况，且受到“非理性”投资者一时兴致或一时爱好的影响，那么这种联系就被中断了。异常低的股票价格，是对缺乏信息的外来因素（如非理性市场偏见）的一种反映，会对公司权益融资和可行投资项目（进行理性计算后所确定的项目）形成约束。

上述分析也适用于接管。如果股票市场是近视的，也就是说，只考虑短期利润和红利的增加，那么担心接管的管理人员会将更多的利润用于红利分配，而不是用于那些利润可观的实际投资，比如说研发支出。这种策略会提升股票价格，通常将其称为“短期主义”。政府对此的可能反应是禁止敌意兼并（如日本）。

与上述观点相反的一种观点（即恶意收购有助于福利改善（对被兼并公司的产出和利润而言））需要假设市场有效且接管能使那些责任心“差”的管理人员被更换掉。在这种情形中，恶意竞标者认为，“不负责任”的管理导致了公司股票价格的下滑。恶意竞标者支付的价格大于公司既有的股价。在对“不负责任”的管理人员进行更换并对公司进行重组后，紧接着的利润增加足以弥补购并公司所支付的高价。如果兼并确实具有协同利润，那么兼并就为合并后公司的股东提供了额外收益。

20 世纪 60 年代和 70 年代兴起了集团企业合并的浪潮，紧接着的 80 年代是杠杆收购（leverage buyout）和集团企业解散（如“资产剥离”）。集团企业合并在有些时候的判断基础是，“公司 *A*”获得非关联公司有助于股东降低持有“*A* 公司”股票的风险，因为“企业集团”包含了由公司组成的分散化投资组合。个体通过改变股票投资组合很容易就可以实现分散化操作，所以企业集团合并的上述原因并不成立。（当然，如果由于某种原因，风险规避型个体无法对所持有的股票实现分散化处理，那么上述原因成立的可能性就要大一些。）

如果股价确实反映了公司的基本面，如果认为频繁发布的“消息”对公司未来绩效有着实质性影响，那么即使市场有效，人们还是预期能观察到股票价格的高波动率特征。然而，如果市场无效，长期价格受到“非理性摇摆”的约束，那么股票价格的波动率比有效市场假说下所预测的波动率还要高。金融机构拥有足够多的资源（储备）来对这种波动性进行预测，这似乎是更为合理的情形，它为通常的资本充足率规则适用于金融机构的市场风险提供了一个论据（如在巴赛尔市场风险条例下）。如果还存在系统风险，那么原则上需要政府采取行动来保证资本水平是系统风险边际社会成本的反映而不是系统风险边际私人成本的反映（对任何单个金融机构而言）。系统风险也支持中央银行采取救助金融机构等形式的干预，否则就会促成其他银行的破产（如长期资本管理公司（LTCM），联邦储备制度理事会于 1998 年组织美国银行财团对其进行过救助）。

股票市场或债券市场的有效性对公司金融中的问题有什么含义呢？如果市场有效，那么因期望“融资条件会改善”而延迟实际投资项目的做法没有任何意义：在 EMH 下，当前价格是正确的价格，它反映了项目预期的未来收益。此外，在 EMH 下，通过改变债务融资和权益融资的组合方式来降低公司资本成本是不可能的。Modigliani - Miller 定理（无税收和破产的情形）建议，在有效市场上，资本成本独立于资本结构（即债务 - 权益比率——参见 Cuthbertson 和 Nitzsche（2001a））。资本组合问题也适用于债务期限结构。因为长期和短期公司债券的利

率完全反映了可获得的信息，长期债务和短期债务也不会改变公司的资本成本。例如，在理性预期假设下，低长期利率和高短期当前利率简单地反映了较低远期短期利率的预期。因此，从事前来看，相对于一系列短期债券的“滚动发行”而言，发行长期债券来为投资项目进行融资没有任何优势。

根据EMH下的上述结论，财务主管作为“主动型”管理人员，其作用体现在要么在不同金融资源适当“混合”的方式之间进行选择，要么对新股发行或债券发行的最优时间进行分析，这些作用都是没有意义的。当然，如果市场无效率，那么财务主管通过所选择的红利政策或股票回购计划等就可以改变公司的股票市场定价。

正如人们所想象的那样，经济学家发现很难进行评价的问题是，如果市场不是非常的有效，那么公共政策和公司行为的准确含义是什么（即所谓的“次优”政策）。如果市场有效，那么就有这样一个假定：政府干预是不必要的。如果市场无效，那么政府干预就有了表面理由。然而，在政府政策对代表性投资者行为影响的不确定性给定的条件下，只有当政府感觉到政策的收益超过风险时，政府才应进行干预。市场无效的任何模型需要探知的是，平均来看市场与“效率市场”之间有多大差距，总体而言政府政策对经济福利有什么含义。这是一个相当艰难的任务。

3.3　数学期望、鞅与公平赌博

可以用多种方式来对EMH进行正规表述。我们不希望无谓地涉及这些表述之间的出色之处，因为我们关心的是如何对假设进行检验，如何使用假设来理解资产价格和收益率的行为。然而，我们需要一些形式的定义。这一节首先讨论条件数学期望的一些性质，然后叙述理性预期的一些基本公理，如无偏性、正交性和预测的链式法则等，最后引入鞅和公平赌博的概念。因此，我们就有了一些基本工具来对EMH的各种表述和检验进行考察。

数学期望

如果X是一个随机变量（如英国男士的高度），以概率π_i离散地取值X_1，X_2，X_3，…，那么X数学期望（记为EX）的定义为：

$$\mathrm{E}(X) = \sum_{i=1}^{\infty} \pi_i X_i \tag{8}$$

如果X是连续型随机变量（$-\infty < X < \infty$），具有连续的概率分布$f(X)$（如正态分布），那么

$$\mathrm{E}X = \int_{-\infty}^{\infty} Xf(X)\,\mathrm{d}X \tag{9}$$

RE文献广泛使用条件概率分布。例如，在抛掷骰子的公平赌博中，着陆数字为1到6的概率都为1/6。假设一位朋友让你知道投掷是有偏差的，在一半的情形中出现数字6，其余数字在另外一半情形中出现的概率相等。以你朋友提供的信息为条件，你会将概率调整为：“6”出现的概率为1/2，其余5个数字出现的概率为1/10。因此，你的条件期望值与无偏投掷情形中的期望值不同，因为伴随概率（或概率密度函数）不同。以信息集合（记为）Ω_t为基础的条件期望定义为：

$$E(X_t \mid \Omega_t) = \int_{-\infty}^{\infty} X_t f(X_t \mid \Omega_t)\,\mathrm{d}X_t \tag{10}$$

$f(X_t \mid \Omega_t)$是条件概率密度函数。条件期望可以视为随机变量X_t基于所有相关信息Ω_t的最优预测。条件预测误差定义为：

$$\varepsilon_{t+1} = X_{t+1} - E(Z_{t+1} \mid \Omega_t) \tag{11}$$

平均来看，它始终为零：

$$E(\varepsilon_{t+1} \mid \Omega_t) = E(X_{t+1} \mid \Omega_t) - E(X_{t+1} \mid \Omega_t) = 0 \tag{12}$$

对（11）进行重新排列：

$$X_{t+1} = E(X_{t+1} \mid \Omega_t) + \varepsilon_{t+1} \tag{13}$$

可以将（13）解释为条件期望是终端价值的无偏预测。条件数学期望的另外一个性质是，预测误差 ε_{t+1} 与 t 时刻及之前的所有信息不相关：

$$E(\varepsilon_{t+1}\Omega_t \mid \Omega_t) = 0 \tag{14}$$

这就是众所周知的条件期望的正交性。（14）为什么成立的直觉是，如果可以用 Ω_t 来降低预测误差 ε_{t+1}，那么就可以用它来改进预测，因此，在对 X_{t+1} 进行预测时并没有使用所有的相关信息。也可以认为：最优条件预测就是一种预测，后续预测误差是不可预测的。

注意，最优预测不需要对 X_{t+1} 进行精确预测。每一个 ε_{t+1} 都可能很大，条件期望 E_tX_{t+1} 可能只能解释 X_{t+1} 实际变化的很小一个部分。不能对最优预测进行改进才是最重要的（即用 Ω_t 来减小预测误差 ε_{t+1}）。同样值得注意的是，我们在（14）中限制的只是预测误差的均值行为。条件预测误差的方差（记为 $E(\varepsilon_t^2 \mid \Omega_t)$）不必是常数，事实上它是部分可预测的。

考虑在一月份（t 时刻）进行的一项预测，你先在二月份（$t+1$）对变量 X 在三月份（即 X_{t+2}）的结果进行预测，然后在一月份对上述结果再进行预测。例如，X_t 就是温度。预测的数学表示是：

$$E_t[E_{t+1}(X_{t+2})] \tag{15}$$

如果 t 时刻的信息集合 Ω_t 已经被有效地使用，那么你今天就不可能对未来预测的变化进行预测，因此

$$E_t[E_{t+1}(X_{t+2})] = E_t(X_{t+2}) \tag{16}$$

这里的 $E_t(X_{t+1})$ 等价于 $E(X_{t+1} \mid \Omega_t)$。这就是迭代期望法则。

我们在这里讨论了无偏性、正交性和迭代期望三个性质，对条件数学期望而言三者都成立（由数学“逻辑”可知）。理性预期要做的是在真实概率分布的基础上假定个体的主观预期等于条件数学期望。因此，假设代表性投资者的行为如下：所形成的主观预期等于真实经济模型的条件数学期望。（这也经常被称为“Muth 理性预期”，Muth（1961）。）

为了获得这个假设所必需的一些直观感觉，考虑一个简单的供给需求模型，比方说，小麦的供给需求模型。供给曲线和需求曲线容易受到随机冲击（如供给方面气候的变化，需求方面如饼干等以小麦为原料的产品“偏好”变化）。从概念来看，“理性”的农场主必须确定每一价格下小麦的供给和其他农场主的小麦预期供给（在诸如技术、投入产品价格等已知的基础上进行预测）。他对影响需求 x_t^d 的因素（如收入等）进行类似计算。然后，将供给冲击和需求冲击的预期值设定为零，他就可以求解预期均衡价格。因此，在（Muth）理性预期下，农场主的行为表明，他们好像使用的是竞争性随机模型。由于供给函数和需求函数的随机冲击，均衡（或预期）价格与最终价格之间的差异是不可预测的随机误差。农场主无法获得其他信息来进一步减小这种误差（即理性预期的正交性成立）。随机简化形式为：

$$P_{t+1} = P_{t+1}^e + \varepsilon_{t+1} = f(x_t^d, x_t^s) + \varepsilon_{t+1} \tag{17}$$

这里的 $P_{t+1}^e = f(x_t^d, x_t^s)$ 是基于已知因素 x_t^i 的均衡价格，它会影响供给和需求。预测误差为随机变量 ε_{t+1}。因此，在理性预期下，经济中的不确定性或随机性（如天气或新产品创新）导致了代表性投资者的预测误差。

为了对代表性投资者的主观预测是否符合条件数学期望公理进行检验，我们要么需要个体主观预期的精确测度，要么需要知道代表性投资者所使用的正确模型形式。对预期进行的调查数据为每个投资者的主观预期提供了一个“带有噪音的”替代变量。如果我们要用第二种方法（即使用正确的经济模型）来对实际预测误差是否具有条件数学期望的性质进行检验，那么我们需要从“经济学家书架”（如凯恩斯主义、货币主义和实际商业周期等）上的模型中选择一个特殊的模型。显然，这些模型的预测误差不遵循理性预期公理可能的原因是，研究人员从“书架”上选出了错误的模型。（也就是说，现实世界中代表性投资者实际上使用的是不

同的模型。）后者为理性预期的支持者提供了一个方便的托辞，因为始终可以这样说，预测误差不遵循公理不是因为代表性投资者“非理性”，而是因为使用了错误的经济模型。

鞅与公平赌博性质

假设随机变量 X_t 具有如下性质：

$$E\ (X_{t+1} \mid \Omega_t)\ = X_t \tag{18}$$

那么称 X_t 为鞅。在（18）给定的条件下，对未来各期 X_{t+j}（$j \geqslant 1$）的最优预测就是现值 X_t。一旦经济人知道了 X_t，Ω_t 中的其他任何信息都无助于预测的改进。随机过程 y_t 是一个公平赌博，如果

$$\mathrm{E}\ (y_{t+1} \mid \Omega_t)\ = 0 \tag{19}$$

因此，在给定 Ω_t 的条件下，公平赌博具有零预期“收益”的性质。由此可以得到如下平凡结论：如果 X_t 为鞅，那么 $y_{t+1} = X_{t+1} - X_t$ 是一个公平赌博。因此，有时也称公平赌博为鞅差分。公平赌博的一个例子是抛掷同质硬币，出现头面就获得 1 美元的偿付，出现反面就获得 -1 美元的偿付。公平赌博性质意味着随机变量的“收益”y_t 平均来看为零，即使代表性投资者在进行预测时已经使用了 Ω_t 中所有可获得的信息。

EMH 的定义包含了无法预测股票收益 $y_{t+1} = R_{t+1} - \mathrm{E}_t R_{t+1}$ 的公平赌博性质，这里的 $\mathrm{E}_t R_{t+1}$ 是一些经济模型所给定的均衡预期收益。公平赌博性质意味着平均来看超额收益为零。因此，投资者在某个特殊时期可能经历了收益的大起大落（相对于均衡预期收益 $\mathrm{E}_t R_{t+1}$ 而言），但这些“赌博”序列的平均结果应为零。如果假设投资者的必要均衡收益为常数（$=k$），那么公平赌博性质意味着：

$$\mathrm{E}\ [\ (R_{t+1} - k)\ \mid \Omega_t]\ = 0 \tag{20}$$

对收益是否违反常数均衡收益假设下公平赌博性质的直接检验是考察历史数据是否对收益有预测效果。假设有如下线性回归方程：

$$R_{t+1} = \alpha + \beta' \Omega_t + \varepsilon_{t+1} \tag{21}$$

如果 $\beta' \neq 0$（或 ε_{t+1} 序列相关），那么公平赌博性质就被违背了。公平赌博性质的检验在这里等价于对理性预期正交性的检验。

Samuelson（1965）指出，常数必要收益率假设下的公平赌博结果可以在投资者偏好的确定性（限制性）假设下来导出。所有投资者都有相同的常数时间偏好率和同质型预期，都属于风险规避类型。因此投资者会持有那些收益率最高的资产，而不考虑风险，从而所有资产的收益都会相等，必要（实际）收益率等于实际利率，这又反过来等于常数时间偏好率。

鞅与随机游走

称随机变量 X_t 是飘移参数为 δ 的随机游走，如果

$$X_{t+1} = \delta + X_t + \varepsilon_{t+1} \tag{22}$$

ε_{t+1} 是独立同分布的随机变量，

$$\mathrm{E}_t \varepsilon_{t+1} = 0$$

$$\mathrm{E}_t\ (\varepsilon_m \varepsilon_s \mid X_t)\ = \begin{cases} \sigma^2, & m = s \\ 0 & m \neq s \end{cases} \tag{23}$$

在无飘移项的随机游走中，$\delta = 0$。显然，X_{t+1} 为鞅，$\Delta X_{t+1} = X_{t+1} - X_t$ 为公平赌博（对 $\delta = 0$ 而言）。因为 ε_t 是独立型随机变量，所以 $m \neq s$ 时的联合密度函数 $f\ (\varepsilon_m,\ \varepsilon_s)\ = f\ (\varepsilon_m)\ f\ (\varepsilon_s)$ 无论是线性函数还是非线性函数都将 ε_s 和 ε_m 之间的任何相关性排除掉了。鞅比随机游走的限制条件更少，因为对鞅而言，只需要 ε_s 和 ε_t 的无关性（即不是线性相关）。此外，随机游走比鞅的限制条件更多是因为鞅不要求高阶矩（如 σ^2）具有统计独立性。例如，如果股票价格（包括红利支付）是鞅过程，那么连续价格变化是无法预测的，但鞅过程允许根据历史方差来预测价格变化的条件方差 $E\ (\varepsilon_{t+1}^2 \mid X_t)$。但是，如果价格遵循随机游走，那么就不允许存在时

变条件方差。

EMH 的规范定义

假设 Ω_t 表示预测收益的（当前和过去）所有相关信息，市场参与者 p 的信息集合记为 Ω_t^p（假设信息获得没有成本）。在有效市场上，假设代表性投资者知道所有相关信息（即 $\Omega_t^p = \Omega_t$），知道收益可能结果的完整（正确）概率密度函数

$$f^p(R_{t+n} \mid \Omega_t^p) = f(R_{t+n} \mid \Omega_t) \tag{24}$$

因此，在 EMH 下，投资者知道产生预期收益的正确经济模型，使用所有相关信息来形成预期收益的“最优”预测。这就是 EMH 的理性预期部分。

代表性投资者有时候会看到，他们的预测存在误差，误差与事后利润或损失有关，

$$\eta_{t+1}^p = R_{t+1} - E^p(R_{t+1} \mid \Omega_t^p) \tag{25}$$

这里的上标 p 表示市场参与者，预测误差以投资者所使用的信息集合为条件。预期收益或均衡收益包括对市场（系统）风险的补偿部分，它使投资者能获得正常利润。（准确地说，风险溢价的决定因素取决于所假设的价值评估模型。）EMH 假设超额收益（或预测误差）只对消息产生反应，因此，相对于可获得的信息而言，η_{t+1}^p 为创新（innovation）（即 RE 的正交性成立）。

实证检验需要对“相关信息”的结构进行定义，著名的信息结构类型有三种。

* 弱式：信息集合只包含历史价格（收益）信息。

* 半强式：信息集合包含所有公开信息（也包含历史价格和收益信息）。

* 强式：价格反映了所有可能的信息，包括“内部信息”（如即将宣布的接管或合并方面的信息）。

在实证研究中，对 EMH 的检验通常考虑半强式。现在，我们将 EMH 的基本思想进行概括。

（ⅰ）所有投资者按照收益的（或确定价格的）均衡模型行事；

（ⅱ）为了确定均衡收益（或基础价值），所有投资者以相同的方式对所有相关信息进行加工。根据预测进行时刻所获得信息来看，预测误差是不可预测的。

（ⅲ）在一个“赌博”序列上投资者不能获得超额利润。

当研究人员所考虑的经济模型就是投资者所使用的正确经济模型时，对（ⅱ）的检验可以简化为对理性预期公理（如无偏性和正交性）的检验，一般称这种检验为信息效率检验。基于（ⅲ）的检验稍微有些不同。超额收益是可预测的，但投资者是否可以获得超额收益取决于用风险和交易成本进行正确调整后的收益。对（ⅲ）最好的表述有可能是 Jensen（1978）。

“相对于信息集合 Ω_t 而言市场有效率，如果基于 Ω_t 的交易不可能获得经济利润。这里的经济利润是指用风险调整后的收益，再减去所有的成本。”

3.4 检验 EMH

这一节对 EMH 的一些检验程序进行概述，将这些检验程序分为如下三种类型是非常有用的：

（ⅰ）检验超额收益 $\eta_{t+1}^p = R_{i,t+1} - E_t^p R_{i,t+1}$ 是否与 t 时刻及之前的信息 Ω_t 相独立。为了对这个命题进行检验，考虑

$$R_{i,t+1} = E_t^p R_{i,t+1} + \gamma' \Omega_t + \omega_{t+1} \tag{26}$$

这里的 $E_t^p R_{i,t+1}$ 等于均衡预期收益。如果信息 Ω_t 的解释能力增加，那么 $R_{i,t+1} - E_t^p R_{i,t+1}$ 就是可预测的。这是对信息效率进行检验，它要求代表性投资者所使用的均衡资产定价模型有显式

表示。

（ii）检验实际“交易规则”（诸如买入小盘股、卖空大盘股等“主动型”策略）在扣除交易成本和对“主动型”策略（系统）风险进行补偿后是否能获得超额利润，超额利润通常以“被动型”策略（如持有标准普尔500指数）为基准来进行度量。这些检验能模拟可能的投资行为，包括对显式交易规则（如价值增长）、基于回归方程的主动策略和所谓的“异像”等进行检验。

（iii）检验市场价格是否始终等于基础价值。这些检验运用历史数据、使用某种形式的红利折现模型（RVF）来计算基础价值（或基础价值的方差），然后检验实际价格的方差是否与基础价值的变化相一致（如Shiller的波动率检验）。

从原理上讲，上述检验并不相互排斥，但它们在实践中会给出不同的推断。事实上，在理性泡沫的特殊情形中，类型（i）的检验即使支持EMH，（根据原理）也会与类型（iii）的检验相矛盾。这是因为，如果市场出现理性泡沫，那么理性预期将会形成，预测误差与Ω_t独立，但价格不等于基础价值。

3.5　使用调查数据

假设我们拥有个体主观预期时间序列的准确调查数据，那么不用选择特殊的模型就能考察这些预测是否违背了理性预期公理。因此，联合假设检验就简化为只对信息效率假设进行检验。无论代表性投资者实际使用的是什么均衡模型，我们所获得的结论都有效。尽管使用调查数据的检验好像可以避免EMH检验领域中的一个关键问题（所使用的均衡模型是什么），这些检验本身还是存在一些内在的困难。

调查数据有时可以通过单个投资者对经济变量（如将来的通货膨胀率、外汇汇率或利率）的预期来获得，可能是所收集的与个体预期有关的定量信息，例如，他可能会回答“明年这个时候的利率为10%”。个体i的这些信息提供了预期值的时间序列$Z^e_{i,t+j}$。利用历史数据，我们能直接计算出每个个体在所有时期上的预测误差$\varepsilon_{i,t+j}=Z_{i,t+j}-Z^e_{i,t+j}$。我们不需要知道个体用来预测$Z_{i,t+j}$的精确模型，到目前为止，可以通过如下回归模型来对信息效率进行检验：

$$Z_{i,t+j}=\beta_0+\beta_1 Z^e_{i,t+j}+\beta_2{}'\Lambda_t+\varepsilon_{i,t+j} \tag{27}$$

检验的零假设为：$H_0:\beta_0=\beta_2=0$且$\beta_1=1$。如果H_0没有遭到拒绝，那么由（27）可知预测误差平均来看为零

$$E\ ((Z_{i,t+j}-Z^e_{i,t+j})\ |\ \Omega_t)\ =E\ (\varepsilon_{i,t+j}\ |\ \Omega_t)\ =0 \tag{28}$$

且预测误差与t时刻能获得的信息Λ_t独立。受限信息集合$\Lambda_t\subset\Omega_t$（在t时刻及之前就已经知道的所有变量（如历史利率和股票收益））。因为预测误差$\varepsilon_{i,t+j}$与t时刻的信息独立，我们也要求$\varepsilon_{i,t+j}$序列不相关（这只适用于$j=1$的情形）。

与预期有关的调查数据通常只能以“总量”形式（即一群个体）获得（如调查的所有参与者在任意时期$t+j$上的平均预测值），很明显，这些结果的解释存在很多问题。例如，如果小样本个体中的某个人所表现出来的行为违背了信息效率假设，那么这可能会导致理性预期公理遭到拒绝。然而，在下述情形中，绝大多数人认为在很大程度上应该支持信息效率假设。

事实上，即使我们拥有个体预期的调查数据，也没有任何激励措施来保证个体显示真实预期，在现实世界中，预期调查数据就是按照这种方式产生的（例如，投资者用大量投资来支持他们的预感）。换句话说，调查数据可能拒绝理性预期下的信息效率假设，因为没有激励措施来保证调查参与者显示真正的预测结果，即使存在预测错误，参与者也不会有任何损失。另外一个问题是，调查的参与者不是市场上实际交易者和“做市商”（即那些“处于边际上”而不是“边际内”的交易者）的典型代表。最后，尽管可以使用一些计量经济技巧（如工具变量估计）来对调查数据的随机误差进行度量，这些方法还是不能解决预期的错误度量问题，

错误度量是个体对真实预期不进行准确报告所造成的系统性错误。

当调查数据是定性数据或分类数据时，这些检验甚至会出现更多问题。在这种情形中，参与者会对如下问题进行回答：“一年中的利率是：（a）更高；（b）更低；（c）与当前利率相同”。我们必须将这种数据“转换成”定量数据，就目前而言，所有可以应用的方法都要求施加一些限制性假设，而这些假设可能导致所考虑的检验无效。

这一领域的应用工作浩如烟海，因为上面所讨论的这些困难，研究方法在过去10年中并没有得到发展。使用调查数据直接对理性预期的无偏性和信息效率进行检验的研究（如Pesaran（1987），Sheffrin（1983），Taylor 1988）倾向于频繁拒绝理性预期公理（也可参见其他读物：Batchelor和Dua（1987），Cavaglia、Verschoor和Wolf（1993），Ito（1990），Frankel和Froot（1988））。就这一点而言，读者可能会觉得，进行理性预期假设是不值得的。如果预期不理性，为什么要继续讨论理性假设下的资产定价模型呢？这个问题的答案是：基于调查数据的检验并不是最后的检验，这些检验存在如上所述的一些局限性。因此，基于市场中实际“交易”收益或价格数据来对理性预期进行直接检验，为基于调查数据的直接检验提供了有益的补充。

正交性质和交叉方程限制

如果无法获得调查数据，那么只有在其他假设下才能对市场有效的零假设进行检验。这个假设是，研究人员所选择的均衡定价模型就是市场参与者实际使用的模型，因而是“正确的”模型。为了用一种简单的方式来说明正交性检验和理性预期下的交叉方程限制，假设Z_{t+1}的均衡定价模型可以表示为

$$\mathrm{E}_t^p Z_{t+1} = \gamma_0 + \gamma' x_t \tag{29}$$

这里的x_t是均衡定价模型所建议的变量集合。在所选择的均衡模型条件中，信息效率（或正交性）检验涉及如下回归方程

$$Z_{t+1} = \gamma_0 + \gamma' x_t + \beta' \Lambda_t + \varepsilon_{t+1} \tag{30}$$

正交性检验的令假设为H_0: $\beta' = 0$，也可以对所选择定价模型建议的（γ_0，γ'）上的限制条件进行检验。对$\beta' = 0$的检验就是对均衡定价模型中决定因素x_t完全解释$E_t Z_{t+1}$的行为进行检验。当然，运用其他均衡定价模型也可以对信息效率进行检验。

注意，在上面讨论的所有检验中，ε_{t+1}必须序列无关（因为$\varepsilon_t \subset \Omega_t$，$\Omega_t$为完全信息集合）。然而，$\varepsilon_{t+1}$不必同方差，$\varepsilon_{t+1}$的方差可能随着时间发生变化，或依赖于其他经济变量，但这没有违反信息效率假设，因为信息效率只依赖于分布的一阶矩——即ε_{t+1}的期望值。

交叉方程限制

对“信息效率”更强的检验涉及交叉方程限制。一个简单的例子足以说明这一点，在后面的章节中，向量自回归模型（Var）要涉及一些复杂的交叉限制，这个例子是对此的有用引言。我们在这里将代数运算减少到最小程度（完整的推导见附录）。考虑仅在$t+1$期末支付不确定性红利的一只股票（这也包括已知时刻$t+1$的股票赎回价值）。“理性”的均衡价格为：

$$P_t = \delta \mathrm{E}_t D_{t+1} \tag{31}$$

δ为常数折现因子。现在假设受限信息集合Λ_t =（D_t，D_{t-1}）上的红利预期生成方程为：

$$D_{t+1} = \gamma_1 D_t + \gamma_2 D_{t-1} + v_{t+1} \tag{32}$$

在理性预期假设下，E（$v_{t+1} \mid \Lambda_t$）=0。现在可以说明，均衡定价模型（31）+所假设的显式预期生成过程（32）+理性预期假设——简而言之，EMH，意味着参数之间存在某种限制。为了看清这一点，在理性预期假设下，由（32）可以看到

$$\mathrm{E}_t D_{t+1} = \gamma_1 D_t + \gamma_2 D_{t-1} \tag{33}$$

代入（31）

$$P_t = \delta\gamma_1 D_t + \delta\gamma_2 D_{t-1} \tag{34}$$

将（34）写成

$$P_t = \pi_1 D_t + \pi_2 D_{t-1} \tag{35}$$

其中，$\pi_1 = \delta\gamma_1$，$\pi_2 = \delta\gamma_2$。P_t 对（D_t，D_{t-1}）的回归方程给出了系数 π_1 和 π_2 的估计值。类似地，回归方程（32）给出了 π_3 和 π_4 的估计值：

$$D_{t+1} = \pi_3 D_t + \pi_4 D_{t-1} + v_{t+1} \tag{36}$$

其中，$\pi_3 = \gamma_1$，$\pi_4 = \gamma_2$。然而，如果（31）和（32）是正确的，那么这意味着（34）成立。因此，根据回归方程（35～36），EMH 意味着：

$$\pi_1/\pi_3 = \pi_2/\pi_4 = \delta \tag{37}$$

（γ_1，γ_2）的值可以由 π_3 和 π_4 及（37）直接得到，δ 可以由 π_1/π_3 或 π_2/π_4 来得到。因此，一般而言，我得到的是 δ 两个不同的值（即系统是“过度识别的”）。

我们有四个参数的估计值（即 π_1 到 π_4），而模型只有三个隐含参数（δ，γ_1，γ_2）。因此，在（37）给定的条件下，各个 π 之间只有一个限制条件，限制条件涉及两个不同方程的待估参数。因此，（37）包含一个非线性交叉方程限制，也是“理性预期 + 定价模型”的一个含义。（注意，在施加这种限制后，我们只得到了一个 δ 值。）

对交叉方程限制进行直观解释是可能的。这些限制仅仅为了保证平均超额利润无法获得，确保预测红利的误差与 t 时刻及之前的信息相独立。首先，考虑使用估计方程（35）和（36）所能获得的利润。公平价值的最优预测值由 $V_t = \delta E_t D_{t+1}$ 给出，运用（36）可以得到

$$V_t = \delta\ (\pi_3 D_t + \pi_4 D_{t-1}) \tag{38}$$

一般而言，已经实现的价格与（38）所给出的基础价值不同，因为研究人员能驾驭的信息比代表性投资者要少（即 $\Lambda_t \subset \Omega_t$）。价格由（35）给出，因此利润为：

$$\begin{aligned} P_t - V_t &= (\pi_1 D_t + \pi_2 D_{t-1}) - \delta\ (\pi_3 D_t + \pi_4 D_{t-1}) \\ &= (\pi_1 - \delta\pi_3)\ D_t + (\pi_2 - \delta\pi_4)\ D_{t-1} \end{aligned} \tag{39}$$

因此，对（D_t，D_{t-1}）的所有值而言，只有当

$$\delta = \pi_1/\pi_3 = \pi_2/\pi_4$$

成立时利润才为零，但这个 δ 值施加了交叉方程限制（37）。现在考虑预测红利的误差：

$$D_{t+1} - E_t D_{t+1} = (\pi_3 D_t + \pi_4 D_{t-1} + v_{t+1}) - (1/\delta)\ P_t \tag{40}$$

这里应用了（36）和均衡模型（31）。将（35）中的 P_t 代进去后就得到：

$$D_{t+1} - E_t D_{t+1} = (\pi_3 - \pi_1/\delta)\ D_t + (\pi_4 - \pi_2/\delta)\ D_{t-1} + v_{t-1} \tag{41}$$

因此，如果 $\delta = \pi_1/\pi_3 = \pi_2/\pi_4$，那么预测误差只依赖于 t 时刻的信息（即 D_t 和 D_{t-1}）。对交叉方程限制的检验在 EMH/RE 文献中非常普遍，一般而言，检验比上述简单例子中的情形要复杂得多。然而，无论检验有多么复杂，这些限制只是为了保证平均超额利润无法获得及预测误差与信息集合正交。

上述检验程存在的另外一个问题是，检验以 D_{t+1}生成方程中的特殊数学期望为条件。如果代表性投资者对如何形成预期的表述不正确，那么所估计的参数值（π_3，π_4）很有可能是真实参数的有偏估计。因此，基于这些估计参数的交叉方程限制可能不成立，即使在“正确的模型”中，它们也可能不成立。这就是计量经济学中常见的误设偏差。

对市场效率检验的解释

EMH 假设能零成本地获得信息，市场价格的运动按照实际情形来确定。信息获得、加工及对这些信息做出反应所涉及的时间都没有任何成本，这是一个非常强的假设。如果价格“始终是所有可获得信息的反映”，即使获取信息的成本为零，为什么人们还要投资资源来获取信息呢？显然，任何进行这样操作的人所获得的收益都比那些以零成本观察当前价格的人所获得的收益低，后者能在 EMH 下获得所有相关信息。正如 Grossman 和 Stiglitz（1980）所指出的那样，如果信息的获取有成本，那么价格就不可能完全反映所获取的信息。Stiglitz（1983）也指出，资本市场在任何时刻都不可能完全有效。投机利润是更快获取信息的结果，是既有信

息和新信息的正确解释。因为相对于信息不灵通的投资者而言，那些“信息灵通的”投资者获得了利润，所以可以预期市场会朝着富有效率的方向运动。在这个过程中，“富有经验的投资者”在实际价格高于基础价格时卖出股票，这使得价格更加接近于基础价值。然而，这个过程可能要花很长一段时间，如果代表性投资者对生成基础价值（如红利）的正确模型不确定，那么这个过程会尤其漫长。此外，非理性或“噪声”交易者可能会出现，因此理性交易者必须考虑噪声交易者的行为。因此，在多时期模型中，价格偏离基础价值是可能的。最近，许多研究考察的是代表性投资者的学习机制、序惯交易本质和噪声交易行为。后面的章节将讨论这些问题。

3.6 小结

我们已经考虑了 EMH 的一些基本思想，兼顾了直觉和数学描述，主要结论如下：

* EMH 在对合并接管的合意性、短期主义、金融市场和金融机构管制等公共政策评估方面有着非常重要的作用。

* EMH 意味着（预料不到的）收益是一个公平赌博。

* 可以用预期的调查数据来对（理性预期下）信息效率进行检验。

* 一般而言，对 EMH 的检验需要均衡预期收益的一个显式模型。以均衡模型为条件，根据已知数据无法对收益进行预测。

* EMH 假设投资者有效地加工信息，从而交易金融资产不可能获得持久的超额利润。交叉方程限制条件是对信息效率（理性预期下）和零超额利润的检验。

* 在实证研究中，EMH 失灵的原因可能是信息效率失灵（理性预期下），或均衡收益模型的不合适选择，或简单地，EMH 在“现实世界中”不成立。

附录：交叉方程限制

在发展这些交叉方程限制的过程中存在一些非常细微的问题，我们在正文中给出了交叉方程限制的简化叙述，下面给出完整的推导。研究人员不可能拥有市场参与者所能获得的全部信息，$\Lambda_t \subset \Omega_t$，这意味着正文的（34）中存在一个误差项，它反映的是信息集合的差异，即 $\omega_{t+1} = [E_t(D_{t+1} \mid \Omega_t) - E_t(D_{t+1} \mid \Lambda_t)]$。为了看清这一点，注意股票价格由代表性投资者所能获得的全部信息来决定：

$$P_t = \delta \mathrm{E}(D_{t+1} \mid \Omega_t) \tag{A1}$$

计量经济学家使用完全信息集合的一个子集 $\Lambda_t = (D_t, D_{t-1})$ 来对红利进行预测：

$$D_{t+1} = \gamma_1 D_t + \gamma_2 D_{t-1} + v_{t+1} \tag{A2}$$

接下来使用一个恒等式：

$$\begin{aligned} \mathrm{E}(D_{t+1} \mid \Omega_t) &= \mathrm{E}(D_{t+1} \mid \Lambda_t) + \{\mathrm{E}(D_{t+1} \mid \Omega_t) - \mathrm{E}(D_{t+1} \mid \Lambda_t)\} \\ &= \mathrm{E}(D_{t+1} \mid \Lambda_t) + \omega_{t+1} \end{aligned} \tag{A3}$$

其中，

$$\begin{aligned} \omega_{t+1} &= [D_{t+1} - \mathrm{E}(D_{t+1} \mid \Lambda_t)] - [D_{t+1} - \mathrm{E}(D_{t+1} \mid \Omega_t)] \\ &= v_{t+1} - \eta_{t+1} \end{aligned} \tag{A4}$$

η_{t+1}是使用完整信息集合 Ω_t 时代表性投资者的理性预测误差。注意，$\mathrm{E}(\omega_{t+1} \mid \Lambda_t) = 0$。为了导出（34）的正确表达式，应用（A1）、（A2）和（A3）就得到：

$$\begin{aligned} P_t &= \delta \mathrm{E}_t(D_{t+1} \mid \Omega_t) \\ &= \delta \mathrm{E}_t(D_{t+1} \mid \Lambda_t) + \delta\omega_{t+1} \\ &= \delta\gamma_1 D_t + \delta\gamma_2 D_{t-1} + \varepsilon_{t+1} \\ &= \pi_1 D_t + \pi_2 D_{t-1} + \varepsilon_{t+1} \end{aligned} \tag{A5}$$

其中，$\varepsilon_{t+1} = \delta\omega_{t+1}$。正文中“无利润”条件（39）的完整推导是通过代表性投资者所使用的完整信息集合来确定 V_t：

$$V_t = \delta \mathrm{E}(D_{t+1} \mid \Omega_t) = \delta \mathrm{E}(D_{t+1} \mid \Lambda_t) + \delta\omega_{t+1} \tag{A6}$$

因此，运用（A5）和（A6）有：

$$\begin{aligned} P_t - V_t &= (\pi_1 D_t + \pi_2 D_{t-1} + \delta\omega_{t+1}) - \delta(\pi_3 D_t + \pi_4 D_{t-1} + \omega_{t+1}) \\ &= (\pi_1 - \delta\pi_3) D_t + (\pi_2 - \delta\pi_4) D_{t-1} \end{aligned} \tag{A7}$$

这就是正文中的方程（39）。红利预测是基于代表性投资者所能获得的完整信息（尽管计量经济学家没有这样多的信息），运用（A2），（A1）可由下式给出：

$$D_{t+1} - \mathrm{E}(D_{t+1} \mid \Omega_t) = (\pi_3 D_t + \pi_4 D_{t-1} + v_{t+1}) - (1/\delta) P_t \tag{A8}$$

然而，将（A5）中的 P_t 代入并注意到 $\omega_{t+1} = v_{t+1} - \eta_{t+1}$，我们有

$$D_{t+1} - \mathrm{E}(D_{t+1} \mid \Omega_t) = (\pi_3 - \pi_1/\delta) D_t + (\pi_4 - \pi_2/\delta) D_{t-1} + \eta_{t+1} \tag{A9}$$

因此，上述方程（A9），而不是正文中的方程（41），是正确的表达式。但正文中（41）的推导没有这么复杂，还为我们提供了一些直觉。

第4章

股票收益可预测吗？

目的

- 考察股票、债券和国库券在过去一个世纪中的收益。
- 提供股票收益在不同时期是否可预测的一组检验方法。
- 引入单变量检验，包括收益相关图（自相关性）和方差比统计量，引入协整和误差修正模型（ECM）等多变量检验。
- 使用蒙特卡罗模拟（MCS）对单变量检验的“功效”进行评价。
- 对股票收益适用的非线性模型和区制转换模型（regime switching model）进行评价。
- 在回归方程预测的基础上，考察“主动型”交易策略（择时交易）的赢利性。

4.1 一个世纪的收益

图1是美国从1915年1月到2004年4月的月度实际标准普尔500指数的图形。简单地说，实际指数就是名义指数除以总的价格指数（如批发或消费价格指数），它反映的是持有分散化股票投资组合的（商品和劳务）购买力变化，这个投资组合能模拟标准普尔500指数。20世纪30年代的大萧条和2000—2003年的股灾清晰可见，股票指数在20世纪90年代的长期上涨也非常明显。股票指数是非平稳序列（或一阶自积序列，I（1）），它的均值不是常数，而是随着时间趋于上升。图2中指数的月度收益（不包括红利收益）似乎是平稳的（即I(0)），（无条件）均值和方差为常数。图2中较大的“异常值”（即绝对值非常大的收益）意味着无条件收益的分布是具有肥尾的非正态分布，分布可能不是对称分布，因为大的负收益出现频率比正收益高，这也意味着分布不是正态分布。收益的柱状图可以对此进行证实，肥尾和向负方向偏斜的特征清晰可见（图3）。

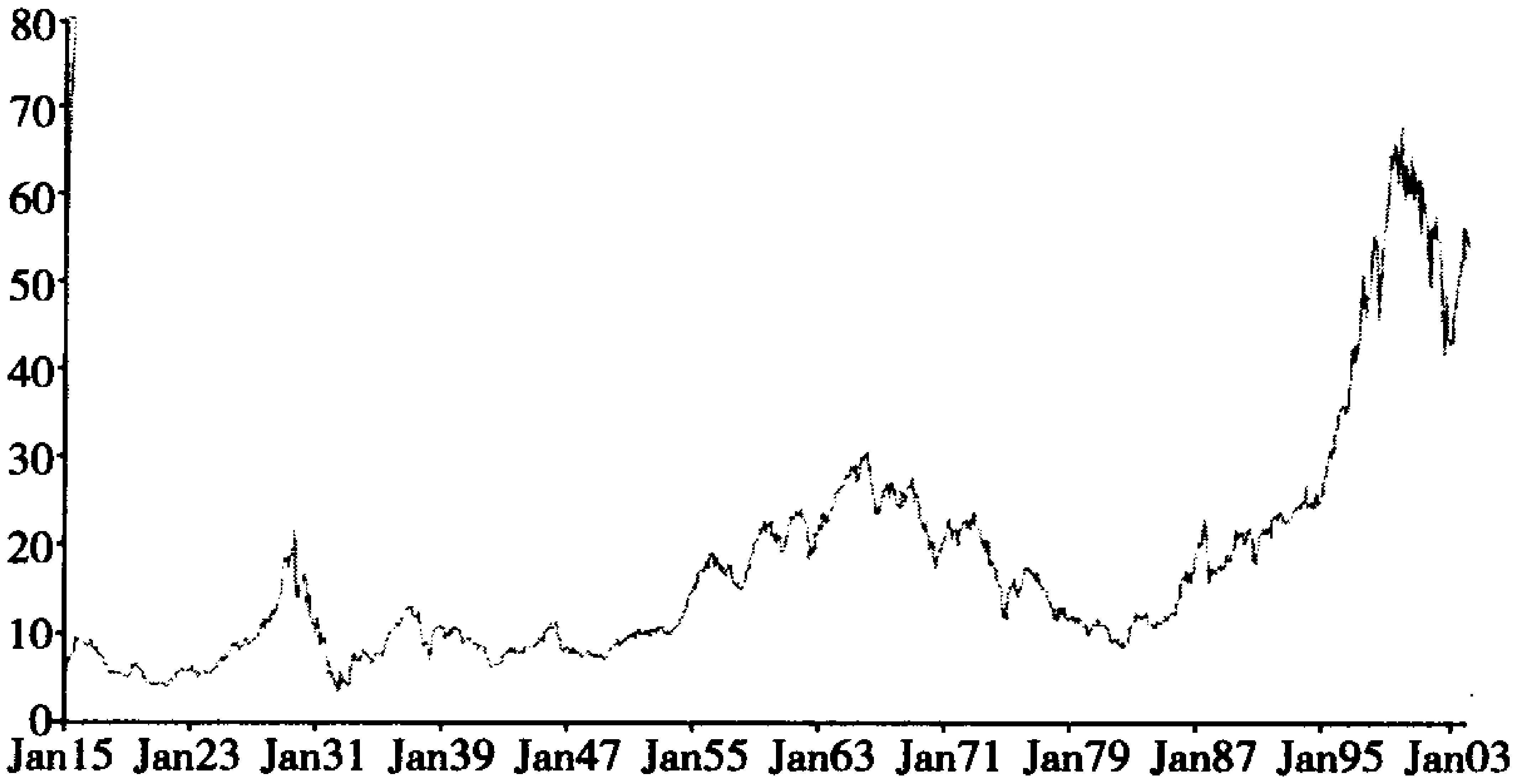


图 1　美国实际的标准普尔 500 指数（1915 年 1 月～2004 年 4 月）

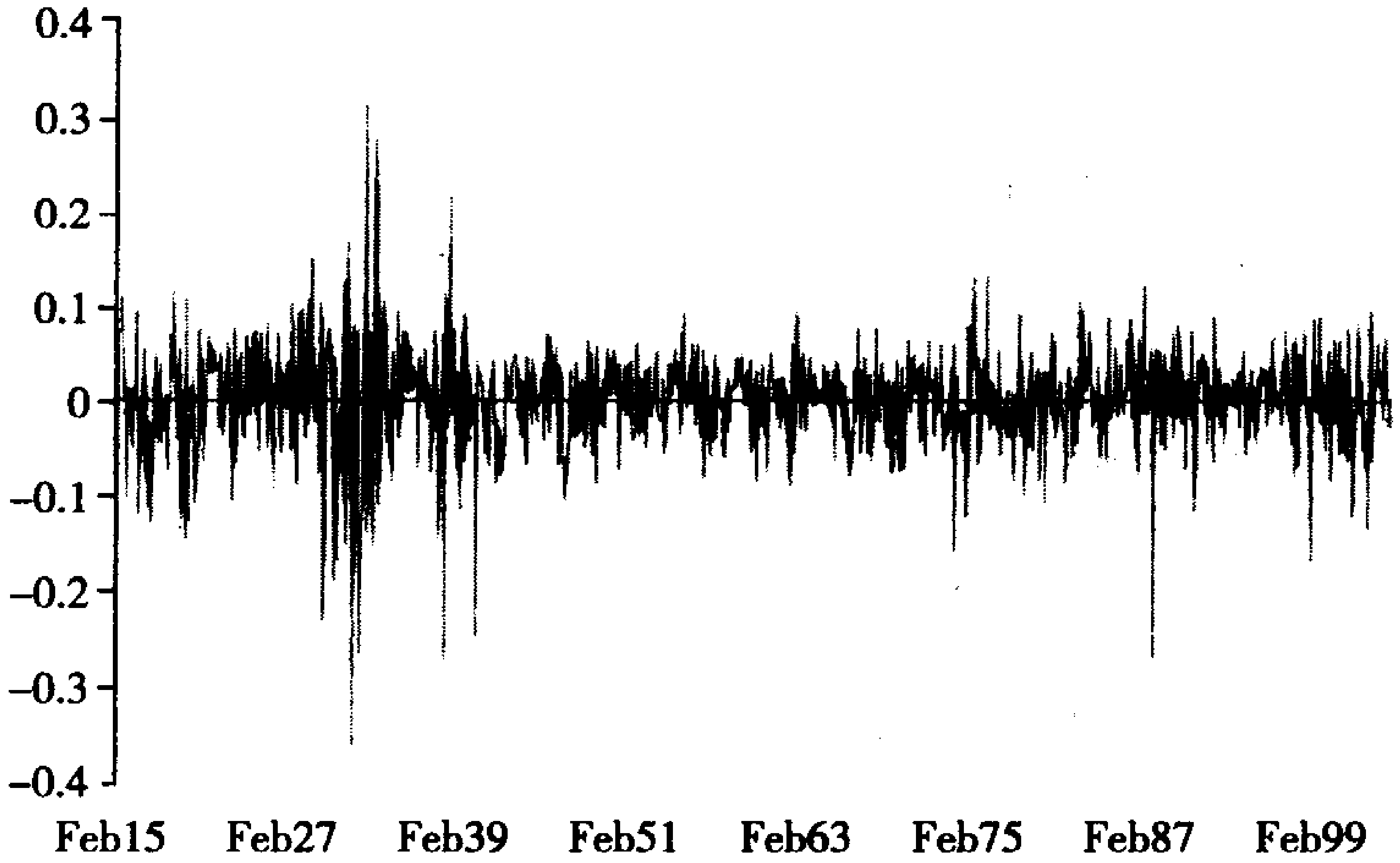


图 2　美国标准普尔 500 指数的实际收益（月度，1915 年 2 月～2004 年 4 月）

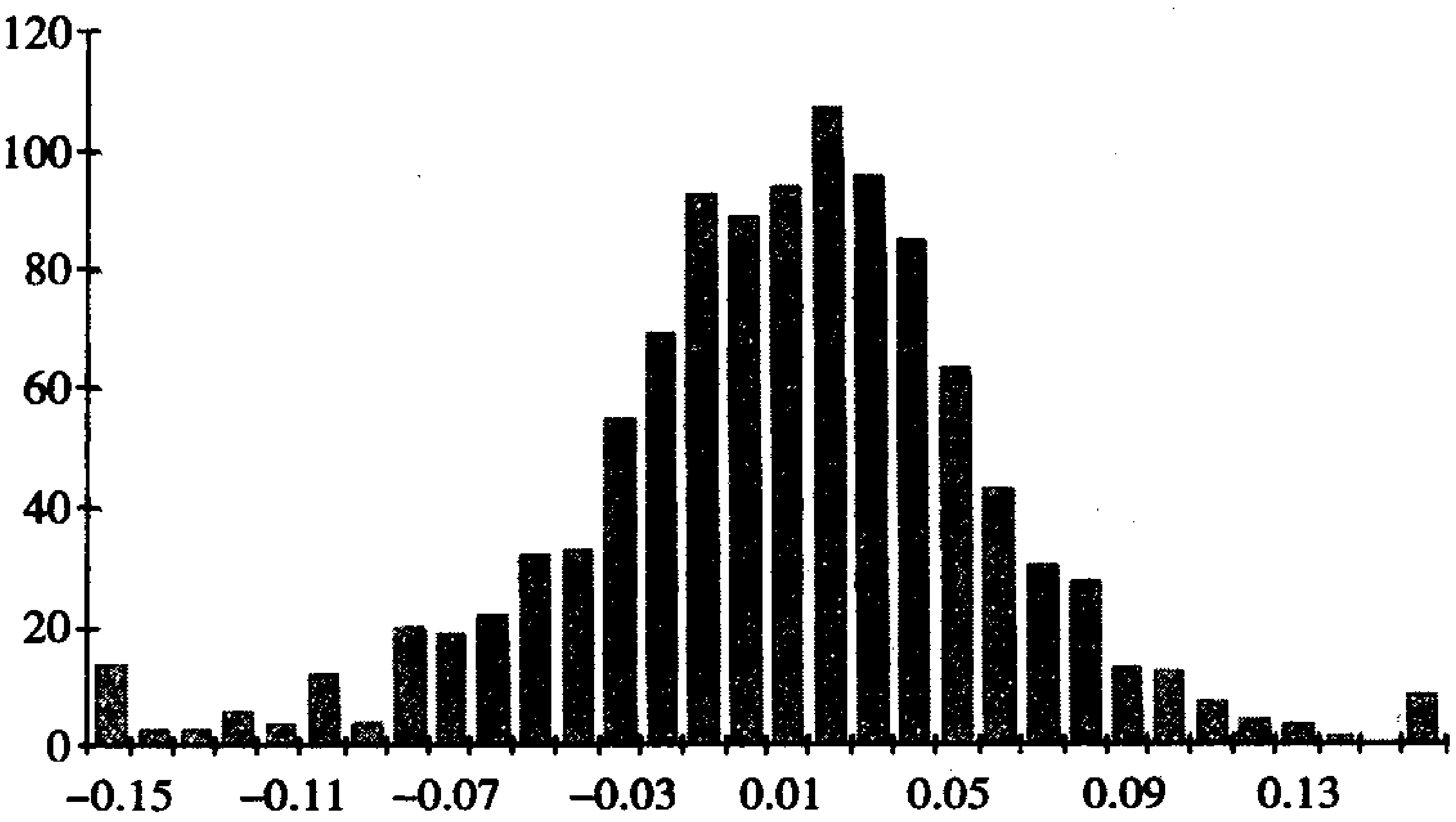


图 3　美国标准普尔 500 指数的实际收益（月度，1915 年 2 月～2004 年 4 月）

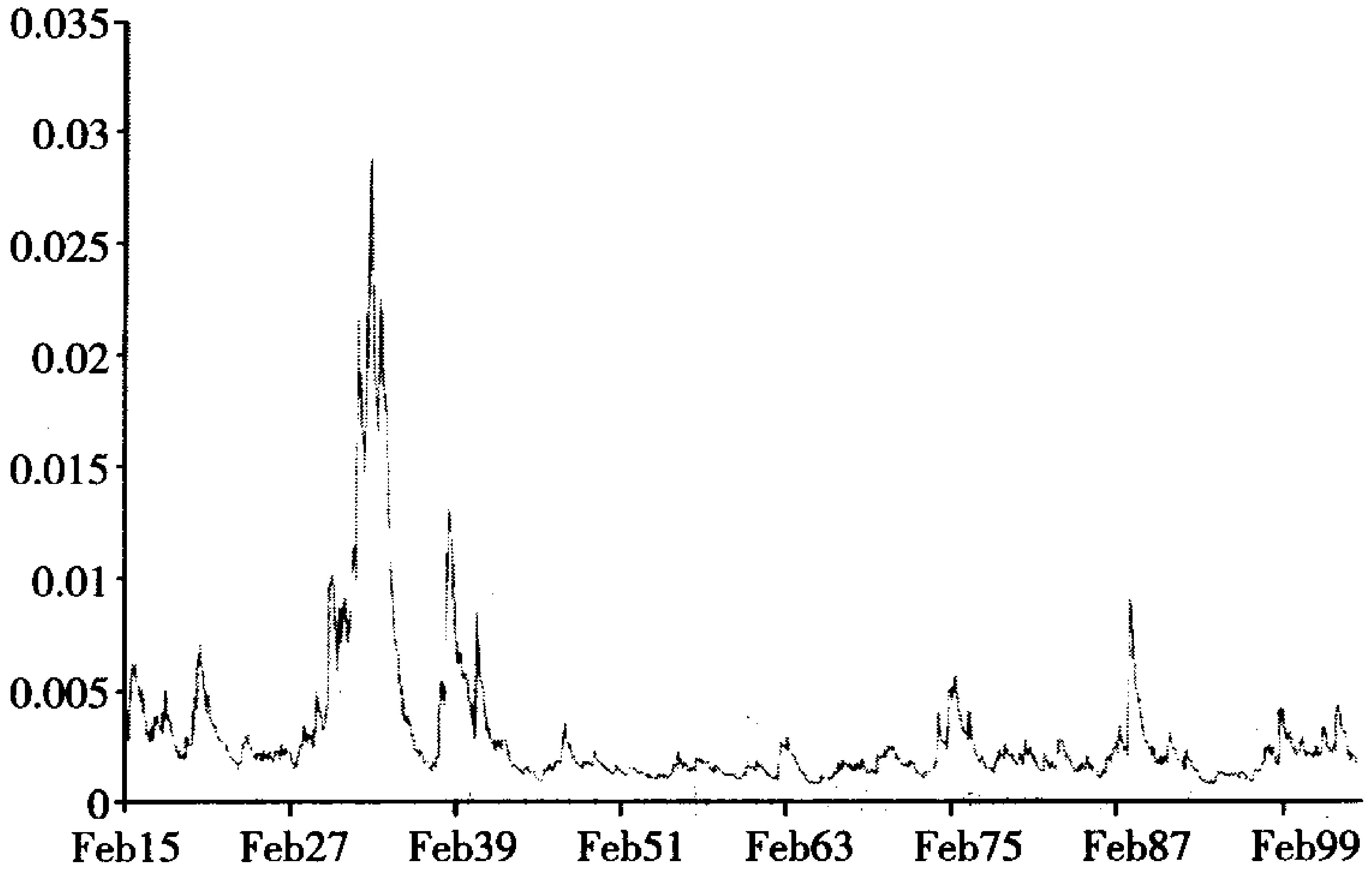


图4　条件方差，GARCH（1，1）模型（美国，1915年2月~2004年4月）

我们转到图2，显然，月度收益的波动率经历了平静时期（如20世纪五六十年代）和动荡时期（如20世纪30年代、70年代和世纪之交时期）。一旦收益的波动性变得很高，收益就会在一段时间上处于高波动状态中；类似地，当收益相对较小时（正或负），收益倾向于在一段时间上处于相对较小的状态。因此，波动率具有条件自回归性质。因为波动率在计量经济学家眼中以异方差而闻名，所以图2中波动率的行为被称为是遵循自回归条件异方差（autoregressive conditional heteroscedasticity，ARCH）过程。这类模型的一个特殊例子是所谓的GARCH（1，1）模型（参见第29章），当用它来拟合图2中1915年2月~2004年4月期间上的数据时，我们就得到如下结果：

$$R_{t+1} = 0.003\ 15 + \varepsilon_{t+1} \qquad \varepsilon_{t+1} \mid \Omega_t \sim N(0,\ h_{t+1})$$

$$[2.09]$$

$$h_{t+1} = 0.000\ 71 + 0.8791 h_t + 0.0967 \varepsilon_t^2$$

$$[2.21]\quad [33.0]\qquad [4.45]$$

（括号中的数字是 t 统计量）

这里的 h_{t+1} 是收益的条件方差。平均（实际）收益为每月0.315%（每年3.85%）。条件方差 h_{t+1} 的GARCH方程是股票收益的一个典型结果，$t+1$ 时的条件方差 h_{t+1} 以 t 时的波动率 h_t 和“预料不到”收益的平方 ε_t^2 为条件。滞后项 h_t 相对较大的系数0.8791意味着，如果波动率高（低），那么收益在一段时期中的波动率都高（低）；也就是说，它具有自回归性质。无条件波动率为

$$\sigma^2 = 0.000\ 71 / (1 - 0.8791 - 0.0967) = 0.000\ 727\ 6$$

这说明标准差为0.026 97（每月2.697%）。图4是上述方程中所使用的时间序列 h_{t+1}，从图形上可以明显看出波动率的持续性。值得注意的是，在每日收益、每周收益和每月收益中都能发现条件波动率的这种GARCH类持续效应，但低频数据（如时期长度超过三个月的收益）通常不会出现这种GARCH效应（但Engle、Lilien和Robins（1987）给出了一个季度期限结构数据中存在ARCH效应的例子）。换句话说，短期收益的波动率具有持续性，而长期收益波动率不具有这种持续性。

现在来看看股票、债券和国库券的年平均收益和波动率，我们使用的是1900—2000年期间的长期数据序列（表1，Dimson、Marsh和Staunton（2002））。英国、美国股票和世界指数（包括美国）的年算术平均（实际）收益 $\bar{R}$ 为7.2%~8.7%（表1A）。这些（算术平均）收益的标准差每年约为20%，说明持有股票在任何一个特殊年份中的风险都很高。如果（对数）

收益服从正态分布，那么 $\bar{R} - \bar{R}_g =$ （1/2） σ^2。就美国而言，$\bar{R} - \bar{R}_g = 2.0\%$，（1/2） $\sigma^2 = 2.04\%$，两者很接近，说明年度收益率近似服从对数正态分布。

高股票波动率也意味着平均收益的事后测度对好坏年份非常敏感。在100年的时期中，1年的收益（比平均收益高两倍标准差（即40%））可能使平均收益增加40个基点（每年0.4%）。这说明我们对 $\bar{R}$ 的估计存在不确定性。换句话说，如果收益独立同分布（即同方差且无序列相关），那么收益的标准差为每年

$$\sigma_{\bar{R}} = \sigma/\sqrt{n} = 20/\sqrt{101} = 2\%$$

因此，我们有95%的自信认为均值（美国）会落在 $\bar{R} \pm 2\sigma_{\bar{R}} = 8.7 \pm 4 = \{4.7, 12.7\}$ 的范围内，对均值收益的可能结果而言，这是一个非常宽的区域。当然，在任意抽取的任何年度中，实际收益的标准差约为每年20%，表1说明收益的最大下降为57%（1974年的英国），同时也有96.7%的上升（1975的英国）。美国股票市场每单位风险的平均收益 $\bar{R}/\sigma$ 为0.43，夏普比（即每单位风险的超额收益）$SR =$ （$\bar{R} - \bar{r}$） $/\sigma$ 约为0.5。

表1 实际收益：1900—2000

Panel A：实际股票收益（%，年）

	通货膨胀率		实际收益					
	算术平均	几何平均	算术平均	标准差	标准误差	几何平均	最低收益	最高收益
英国	4.3	4.1	7.6	20.0	2.0	5.8	−57（1974）	+97（1975）
美国	3.3	3.2	8.7	20.2	2.0	6.7	−38（1931）	+57（1933）
世界指数（包括美国）	n. a	n. a.	7.2	17.0	1.7	6.8	n. a.	n. a.

Panel B：实际债券收益（%，年）

	通货膨胀率		实际收益			
	算术平均	几何平均	算术平均	标准差	标准误差	几何平均
英国	4.3	4.1	2.3	14.5	1.4	*n. a.*
美国	3.3	3.2	2.1	10.0	1.0	1.6
世界指数（包括美国）	n. a	n. a.	1.7	10.3	1.0	1.2

Panel C：实际国库券收益（%，年）

	通货膨胀率		实际收益		
	算术平均	几何平均	算术平均	标准差	标准误差
英国	4.3	4.1	1.2	6.6	0.7
美国	3.3	3.2	1.0	4.7	0.5

注：

（1）年平均值是1900—2000年期间的平均值。“世界”由16个发达国家组成，包括美国、加拿大、南非、澳大利亚、日本和欧洲国家。

（2）实际收益 = $(R_t - \pi_t)/(1 + \pi_t)$，其中 R_t = 名义收益，π_t = 通货膨胀率。因此，实际平均收益不等于名义平均收益减去平均通货膨胀率（后者是一种近似，只有当通货膨胀率很低时，这种近似才有效）。

（3）标准误差用 $\sigma\sqrt{T}$来进行计算。

数据来源于 Dimson、Marsh 和 Staunton（2002）。

表 2　　股权溢价（%，年）：1900—2000

	相对于国库券			相对于债券	
	算术平均	几何平均	标准差	算术平均	几何平均
英国	6.5	4.8	2.0	5.6	4.4
美国	7.7	5.8	2.0	7.0	5.0
世界指数（包括美国）	6.2	4.9	1.6	5.6	4.6

注：见表 1。

政府长期债券的平均（实际）收益（表 1B）每年约为 2%，每年的波动率约为 10% ~ 14%，比股票的波动率低。美国每单位风险的长期债券收益为 0.21，比股票低得多。国库券（表 1*C*）的平均（实际）收益比债券还要低，每年约为 1%，波动率比债券更低，每年约为 $\sigma \approx 5\% \sim 6\%$。Dimson、Marsh 和 Staunton（2002）对 14 个国家股票、债券和国库券的相对收益进行了考察，结论与美国和英国的上述结论相同。

股权溢价是股票超过债券或国库券的超额收益（表 2）。就美国而言，相对于国库券的算术平均股权溢价比相对于债券的股权溢价要高，数字分别为每年 7.7%（相对于国库券）和 7%（相对于债券）。平均股权溢价的标准误差每年为 $\sigma_R = 20\%$，因此，平均股权溢价的度量有着很大的误差，这反映了前面所注意到的股票收益的高波动率特征。美国股权溢价算术平均值 95% 的置信区间为 7.7% ±4%，这再次说明该区间是一个非常宽的区域。

Dimson、Marsh 和 Staunton（2002）所计算的算术股权风险溢价约为 1.5%，比一些早期研究（Ibbotson Associates（2001），Barclays Capital（2003），Goetzmann 和 Jorion（1999））的结果高。这是因为考察的时期和国家不同，计算方法也不同（如红利和价格变化是否包含在所有国家收益的计算中）。

风险、收益和夏普比

显然，如果私人或国家养老金打算对股票等风险资产进行投资，那么他们需要仔细评估所能获得的收益和所承担的风险。美国的高平均股票收益并不只是因为美国经济生产率很高或美国消费者没有耐心（时间偏好），债券的实际收益也很高。政府指数连动债券（government index - linked bond）的回报不是很高（实际年收益约为 3%）。股票收益高的含义是股票市场的风险回报高。正如我们在上面所看到的那样，这种风险是实际存在的。就美国的 $\bar{R}$ 言，年波动率的范围是 4.7% ~ 12.7%（95% 的置信水平）。美国 7.7% 的股权溢价（相对于国库券）是持有股票市场风险的回报。

利用 CRSP 数据库，图 5 证明了金融第一定律（the first law of finance）：高平均收益意味着高风险（标准差）。对图 5 中按照规模进行排序的十分之一股票组合而言，事后夏普比（$\bar{R} - \bar{r}/\sigma$）（每单位风险的超额收益）约为 0.5。显然，获得高平均收益是一件很容易的事情，简单地承担更高的风险就可以了。图 5 也说明，第二次世界大战后美国政府债券和公司债券的夏普比非常低，约为 0.1（这是建议将全国社保基金的部分资金转移到股票上去的部分理由）。

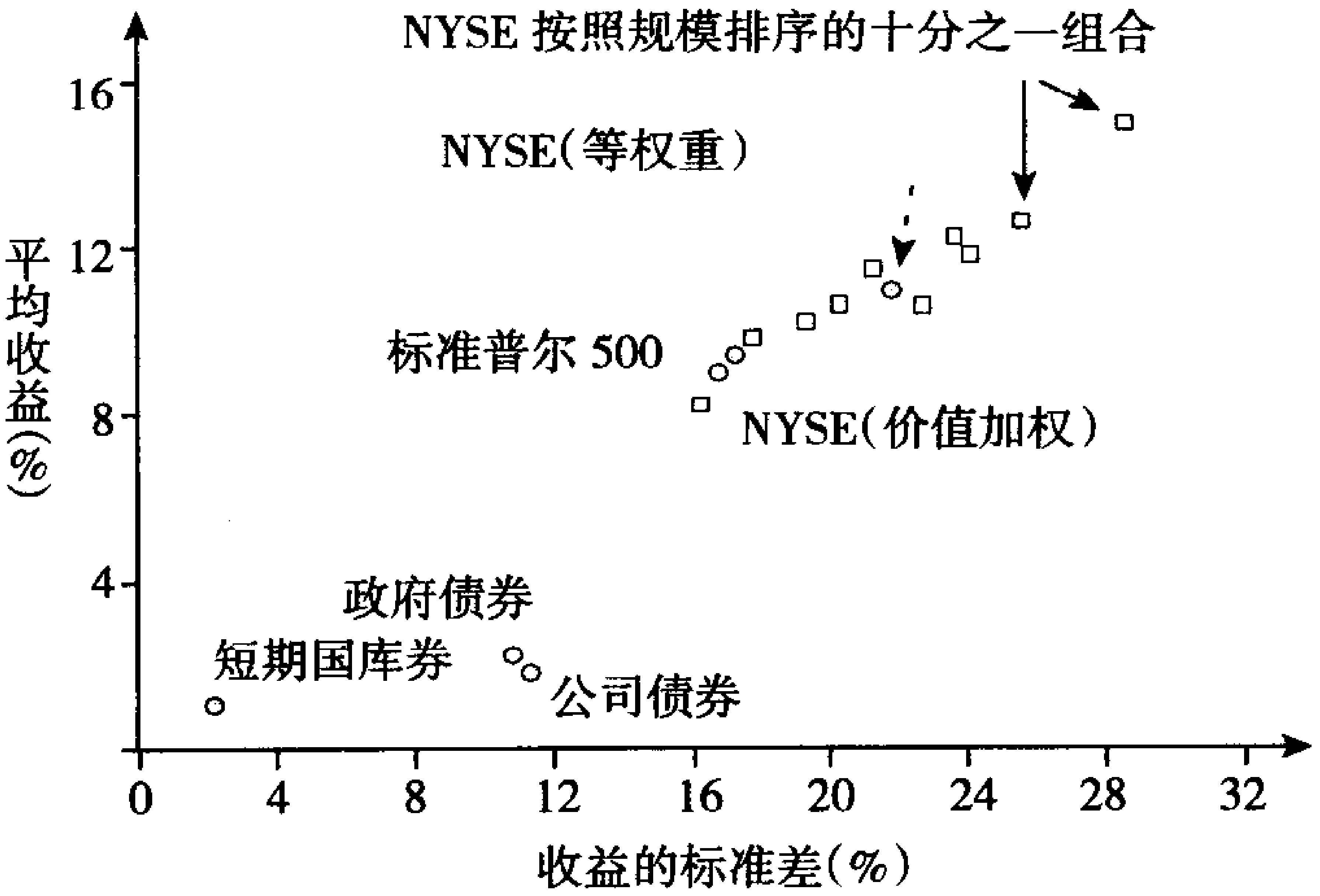


图5　均值和标准差：年平均值，美国实际收益（1947 年后）

利用夏普比而不是平均收益对不同类别的资产进行比较要更好一些，因为夏普比“对风险进行了调整”，杠杆作用对它没有影响。例如，假设当前的股票价格为100美元，平均收益为$\bar{R}=10\%$，标准差为20%，$r=2\%$。如果你将100美元资金全部投资到股票上，那么超额收益为8%，夏普比为0.4。如果你以$r=2\%$的利率借入10%的资金（10美元），用借到的自己和你自己的90%的资金（90美元）去购买股票，那么你的杠杆股票组合的收益为$R_p=(100/90)R\approx 11\%$（通过借钱买股票把你的（平均）收益给“抬”了起来）。如果在风险资产的投资中你借的现金越多，自有资金越少，那么杠杆（预期）收益就越大。扣除利息成本后的收益为

$$R_p \approx 1.1R - 0.1r$$

杠杆组合的超额收益和标准差为（记住$\sigma_r=0$）

$$R_p - r = 1.1\ (R-r)$$

$$\sigma_p = 1.1\sigma_R$$

杠杆组合的夏普比为

$$SR_p = \frac{R_p - r}{\sigma_p} = \frac{R-r}{\sigma_R}$$

它等于无杠杆组合的夏普比，因此杠杆作用对夏普比没有影响。

股权溢价的重要性

对投资者而言，前瞻的股权溢价很重要，因为它为投资股票市场的任何资金相对于无风险收益的未来价值提供了一个估计。对（部分）资产投资于股票的养老基金而言，股权溢价就显得尤其重要。根据养老基金缴纳计划的规定，个体储蓄者在退休时获得“一笔资金”，他通常将这笔资金（所有或一部分）投资到年金上去。“这笔资金”的股票部分对股票平均收益极其敏感。例如，就美国而言，假设年$\bar{R}_g=6.7\%$，0时刻100美元投资在30年后增值到$100(1.067)^{30}=700$美元，如果平均股票收益低一个标准差（即年$\bar{R}_g=4.7\%$），那么100美元的初始投资在30年后只能增值到400美元。这些都是实际收益，因此较高的股票收益使财富增值到7倍，而较低的股票收益只能使财富增值到4倍。对福利养老基金而言（即养老金以最后的工资为基础），前瞻的股票收益对养老基金的偿还能力至关重要。养老基金的负债始终存在，他在未来进行养老金支付，合同是以名义的形式签订的。显然，负债有时是不确定的（如取决于实际收益的增长、退休模式和寿命等）。为了偿还负债，合理的养老金缴纳计划是个体和雇员都对基金做出贡献（以薪水为基础）。最大的不确定性在于，如果投资股票，这些

缴纳的资金在许多年后能获得多少（实际）收益。在经历了2000—2003年的股市衰退后，许多养老基金的托管人认为，股票的未来收益不足以弥补养老基金未来的负债。这就使得这些养老基金只局限于现有成员，新的雇员不得不加入福利计划（因此，股票市场的风险由雇员分担而不是由雇主和雇员一起分担）。此外，那些福利计划赤字的公司（由于股票收益低）不得不增加雇主或雇员或两者的缴纳金额。福利计划对新成员的关闭和一些基金赤字的存在已经成为一个很重要的问题。对英国而言，在经历了2000—2003年股市衰退后这一问题尤其突出。

就世界指数而言，历史上的股票实际（几何）收益为6.8%，这说明在过去的100年中，世界指数追踪基金（tracker fund）（即使是美国或英国的基金）的业绩已经很好（即使将1%的年费扣除）。“主动管理型”基金涉及3%～5%的管理费用，即使它的表现仅与“追踪型基金”相同（“平均”主动型基金似乎就是这样，参见第9章），它的历史业绩也已经很好了。然而，如果你认为前瞻的实际股票收益为（比方说）4%（见后面），那么这对主动型基金而言是没有吸引力的，因为它的管理费用高昂。如果你持有实际年收益为3%的指数连动债券，那么你的状况会更好一些（如John Ralfe于2001年将Boots the Chemist's pension fund的全部资金转移到债券市场）。

对公司而言，计算权益资本成本的关键工具是股权溢价（用公司的β进行调整）的前瞻性度量指标。英国和美国历史上的股权溢价每年高达7%（算术平均），β值为1，平均实际年利率为3%（比方说），公司权益资本的（实际）成本为每年10%。这可能是公司决策的必要收益率（投资项目的内部收益率每年必须高于10%）。但是，如果前瞻的股票收益低于10%，那么，公司显然会降低这一必要收益率，否则增值型投资项目可能被否决。

由于市场权利，国家或私人单位经常受制于价格上限（如价格上涨的速度不能超过消费价格指数减去x%）或“管制收益率”（资本收益每年不能超过y%）。在后一种情形中，前瞻的必要收益率是必需的，如果人们预测股票收益会降低，那么管制者就应确保公用单位的资本只能获得一个更低的资本收益（降低价格）。因此，股票收益预测对养老基金这样的机构投资者、投资于主动型或被动型投资基金的个体投资者、公司投资项目的必要收益率都非常重要。

股权风险溢价的预测

不用说就知道股权溢价的预测存在很大误差。Welch（2000，2001）邀请226位金融经济学家对美国未来30年的平均股权溢价进行预测，预测的（算术）平均值为7.1%，标准差约为2%。预测结果比美国101年间7.7%的股权溢价（表1）低一点点，但金融经济学家们所回答的标准差与历史上2%的标准差大致相同（表1）。当金融经济学家们在对未来股权溢价形成自己看法的时候，他们似乎都“锁定”在历史股权溢价上（心理学上的偏差参见第18章）。

Dimson、Marsh和Staunton（2002）对历史风险溢价进行了调整，他们考虑了上世纪90年代末未曾预料的更高收益的影响、国际投资分散化的增加可能降低必要收益的事实、递减的投资风险（如因为冷战结束、多边贸易谈判取得进展等）和边际投资者风险规避程度降低等因素。他们认为美国股权溢价的预测值为5.4%，而英国为3.7%，世界指数为4%（算术平均，几何平均比它们约低1%）。Campell（2001）使用多种方法的预测结果是：美国的实际股票收益预测值为6.5%～7%（算术平均）或5%～5.5%（几何平均），实际利率的预测值为3%～3.5%。这意味着美国股权溢价的预测范围为3%～4%（算术平均）。事后平均股票收益为

$$\bar{R}_t = \overline{\left(\frac{D_t}{P_{t-1}}\right)} + \overline{\left(\frac{\Delta P_t}{P_{t-1}}\right)}$$

与用历史数据对股权溢价进行直接度量类似，我们可以用“价格比率”来对长期预期资本收益进行度量。如果变量（X/P_t）是平稳的（均值反转），那么就长期而言，

$$\overline{(\Delta X_t/X_{t-1})} = \overline{(\Delta P_t/P_{t-1})}$$

由上述方程可知（用实际资本收益来进行替换），变量X为平均资本收益提供了一个估计值。

X_t 最为明显的选择是红利 D_t 或收益 E_t。收益对股票回购来说是不变的，在有红利支付的情形中，许多公司经常使用股票回购手段，使得红利增长受到扭曲。另一方面，收益比红利的波动率更高，这意味着收益均值的估计具有更大的不确定性。红利价格比和收益价格比的波动非常缓慢，似乎是平稳随机变量（见下面）。认为这些比率经过很长时期之后会到达是违反直觉的，因为这意味着要么红利增长是非平稳的，要么预期收益是非平稳的（参见第12章和13章）。

Fama和French（2002）用平均红利增长和平均收益增长作为未来预期收益的估计值。（使用红利增长等价于使用Gorden增长模型，参见第10章。）他们发现，1872—2000年期间使用红利增长模型所得到的（无条件）实际股权溢价为每年3.54%，而使用历史平均股票收益所得到的股权溢价高达每年5.57%。在过去的50年中（1951—2002），红利增长模型给出的股权溢价为每年2.55%，使用平均股票收益的估计值为每年7.43%。对1951—2000年期间而言，使用收益增长模型估计的股权溢价为每年4.32%，比红利增长模型的结果高，但仍然比每年7.43%的历史平均收益低。

Fama和French注意到，红利价格比和收益价格比能对未来红利增长或未来预期收益进行预测（参见第12章和第23章）。但是，红利价格比对红利的未来增长几乎没有任何解释力。因此，他们认为1950—2000年期间逐渐增加的红利价格比主要是因为预期收益发生了变化，从而在1951—2000年期间的高历史股票收益中有很大一个部分是折现率降低所导致的预期资本收益。因此，大家所一致赞同的观点是，股权溢价的预测值比历史平均值低（也可参见Jagannathan、McGatten和Scherbina（2001））。

4.2 简单模型

本章的主要目的是介绍对股票收益可预测性进行检验的一系列方法。如果你愿意，你可以把回归模型的结果视为“数据描述”或“典型事实”。然而，当预期收益为常数时，“可预测性”违背了信息效率假设；如果将交易成本扣除并对风险进行调整后的超额收益具有持续性，那么这就违背了有效市场假说（EMH）。对EMH进行检验需要一个资产收益均衡模型。我们认为风险资产的均衡预期收益包括无风险利率 r_t（如国库券的收益）和风险溢价 rp_t 两个部分：

$$E_t R_{t+1} \equiv r_t + rp_t \tag{1}$$

就目前而言，我们没有对名义变量和实际变量进行区分，方程（1）对这两种情形都成立。方程（1）是一个恒等式，我们需要有一个能刻画风险溢价的经济模型。对EMH的许多（早期）实证检验都假设 rp_t 和 r_t 为常数，所考虑的回归方程为：

$$R_{t+1} = k + \gamma'\Omega_t + \varepsilon_{t+1} \tag{2}$$

这里的 Ω_t 表示 t 时刻所能获得的信息。作为替代，在方程（2）中我们可以使用超额收益 $R_{t+1} - r_t$。对 $\gamma' = 0$ 进行检验为EMH中的“信息效率”提供了证据。回归检验方程可能有所不同，这取决于所假设的信息：

Ⅰ. 历史收益数据 R_{t-j} $(j = 0, 1, 2, \cdots, m)$；

Ⅱ. 历史预测误差数据 ε_{t-j} $(j = 0, 1, 2, \cdots, m)$；

Ⅲ. 红利价格比、收益价格比率、利率等变量的数据。

对Ⅰ和Ⅱ同时进行考察就得到了自回归移动平均（ARMA）模型，例如ARMA（1，1）模型：

$$R_{t+1} = k + \gamma_1 R_t + \varepsilon_{t+1} - \gamma_2 \varepsilon_t \tag{3}$$

如果我们只考虑弱式有效问题，那么考察 R_{t+1} 与 R_{t-j}（$(j = 0, 1, \cdots, m)$）之间的自相关系数是否非零就可以了。EMH对所有持有期都成立：天、周、月甚至许多年。因此，我们可能会发现EMH在一些时期上不成立，而在其他时期上却成立。

假设上述检验表明信息效率准则不成立，那么 t 时刻的信息就有助于预测未来收益。然

而，如果回归方程的标准差大或 R^2 低，那么投资者以该方程预测的结果进行赌博的风险可能非常的高。因此，值得考察的是，在考虑交易成本之后，这种可预测性是否真正允许交易者通过实际交易来获得超额收益。于是，我们有两种方法对 EMH 进行检验，一种方法是对信息效率进行检验，另外一种方法是对获得超额利润的能力（即扣除交易成本并对事前风险进行调整）进行检验。

精明投资者与噪声交易者

在对可预测性的各种检验方法进行详细讨论之前，对市场上非理性交易者或噪声交易者对价格和收益的含义进行简略讨论是值得的。这使我们能以一种相当简单的方式来引入均值反转和更大波动率的概念。假设市场包含特殊类型的噪声交易者，即正反馈交易者，价格上涨之后他们的股票需求会增加。为了简化问题，假设理性交易者或精明投资者相信预期收益为常数，$1+E_tR_{t+1}=k^*$，或等价地：

$$E_t\ [(P_{t+1}+D_{t+1})\ /P_t]\ =k^* \tag{4}$$

如果精明投资者（基础度量交易者）在市场上出现，那么价格只对新信息或消息产生反应。价格变化是随机的，历史收益不能用于未来收益的预测。现在，考虑在市场上引入正反馈交易者。在有关红利的任何好消息发布后，正反馈交易者会购买股票，拉升股票价格使其高于基础价值。如果理性交易者承认这种错误定价，那么他们会卖空价值高估的股票，价格朝基础价值方向运动。因此价格是均值反转的。

我们立刻就可以看出如下两个事实。第一，均值反转价格对基础变量（即有关红利的消息）存在过度反应。第二，价格波动率比根据基础变量变化所预测的波动率要高。因此，与市场上只存在理性交易者时的价格相比，均值反转价格具有更大的波动率。基于 Shiller（1989）、LeRoy 和 Porter（1981）的早期波动率检验以精确的方式对这一更大的波动率进行了度量。

因为正反馈交易者购买股票，所以在很短的时期内，收益序列正相关：正收益之后出现的是更高的正收益。反之，在出现“坏消息”之后，负收益之后出现的是绝对值更大的负收益。但在一个很长的时期上，收益是序列负相关的，因为理性交易者使价格朝基础价值方向运动。因此，当出现正反馈交易者时，短期收益是序列正相关的，而长期收益是序列负相关的。不同时期上的这种序列相关模式意味着，购买近期的“赢家”（winners）倾向于在下一期产生赢家——这就是动量策略。在一个较长的时期上（比如说 3～5 年），负序列相关意味着你应该买低价股——这就是价值增长策略。此外，上述情形还意味着收益很有可能与红利变化和红利价格比序列相关，因此，R_{t+1} 对 $(D/P)_t$ 的回归经常被用来解释市场上是否存在噪声交易者。

我们也能看到为什么正反馈交易者能造成不同时期中收益方差的变化。假设年收益的方差每年为 15%。如果预期收益为常数，从而收益独立同分布，那么在两年之后，收益的方差将变为 30%。然而，由于均值反转，两年收益的方差将小于一年收益方差的两倍。这是因为价格在短期内会超越（overshoot）基础价值而在长期内却不会出现这种情形（参见下面的方差比率）。

对给定的观察路径是否与市场效率或噪声交易者的存在相一致进行推断可能会很困难。这是因为我们是在一个特殊的均衡收益模型基础上来对 EMH 进行检验，如果均衡收益模型是错误的，那么数据将拒绝 EMH。然而，如果预期收益随着时间发生改变，那么可预测性在本质上并不能反驳 EMH，因为实际收益与 t 时刻已知的某些变量相关。在我们得出“可预测性”可以反驳 EMH 的结论之前，需要对各种均衡收益模型进行仔细考察。我们有多种方法来对股票收益进行检验，下面是对这些检验方法的一个说明（一篇非常全面的参考文献是 Campell（2000））。

4.3　单变量检验

在一个较短的时期中，如一天，可以预期均衡收益为一个常数。因此，实际收益可能是每日异常收益的一个很好近似。Fortune（1991）对标准普尔500指数2 700个每日观察值（1980年1月2日—1990年9月21日）基础上的每日收益进行了统计分析。股票收益由 $R_t = \Delta \ln P_t$ 进行度量。一个典型的回归方程是：

$$
\begin{aligned}
R_t = {} & \underset{(2.8)}{0.0007} + \underset{(2.8)}{0.054\varepsilon_{t-1}} - \underset{(1.9)}{0.037\varepsilon_{t-2}} - \underset{(1.0)}{0.019\varepsilon_{t-3}} \\
& - \underset{(2.8)}{0.054\varepsilon_{t-4}} + \underset{(2.7)}{0.051\varepsilon_{t-5}} - \underset{(3.2)}{0.0017WE} \\
& + \underset{(0.2)}{0.0006\mathrm{HOL}} + \underset{(0.82)}{0.0006\mathrm{JAN}} + \varepsilon_t \qquad (5) \\
& R^2 = 0.0119,\ SEE = 0.0108,\ (\cdot) = t\ \text{统计量}
\end{aligned}
$$

如果交易日为星期一，那么变量WE取值为1，否则取值为0。如果交易日之前是节假日，那么变量HOL取值为1，否则取值为0。如果交易日是一月份中的交易日，那么变量JAN取值为1，否则取值为0。统计上显著的唯一虚拟变量（就上世纪80年代的数据而言）与“周末效应”有关，这意味着价格收益与星期一负相关。在针对“总的”标准普尔500指数的上述回归方程中，一月效应并不显著（但对小公司的股票而言，一月效应仍然是非常重要的）。误差序列相关，因为MA（1）、MA（4）、MA（5）在统计上是显著的。因为前期预测误差 ε_{t-j}（在 t 时刻）是已知的，所以这是常数均衡收益零假设下对信息效率准则的违背。在周或月度数据中，MA模式可能不存在，因此预测误差可能是一个白噪声，从而EMH获得了支持。

然而，上述数据并不意味着EMH不成立，EMH的定义是不能持续地获得超额利润。回归方程只能解释每日股票收益变化的1%（$R^2 = 0.01$），因此，潜在的获利套利机会很可能涉及很大的风险。基于（5）击败市场的策略是反复进行如下操作：在星期五卖空股票，在星期一再将这些股票买回来，这样操作可预测的收益为0.001 7%。但是，如果投资组合由模拟标准普尔500指数的25～30只股票组成，那么这将涉及高昂的交易成本，从而这些交易策略的利润会被抵消掉。然而，股票指数期货的使用能减轻这一问题的严重性。因为方程（5）中的系数是样本数据平均值意义下的结果，所有我们有相当的自信认为这些“平均效应”在未来会持续一段时间。当然，如果 WE 的系数仍为负数，那么重复策略将获得利润（忽略交易成本）。但是，在进行这一系列的重复赌博之前，我们至少可以对系数的短暂稳定性进行检验。此外，这可能需要大量的金融资源，持这种交易策略的投资者在前几周可能不是那么的幸运，他们在开始的时候可能会损失一大笔资金。后者对养老基金管理者等机构投资者而言是否是一个问题取决于绩效评估的频率。0.01左右的 R^2 意味着你进行的是51/49的赌博，如果涉及交易成本，这并不是一个很好的策略。为了充分地对EMH下“异常利润”进行检验，我们需要考察“现实世界”的一些投资策略：对投资组合中的一些特殊股票进行交易，考虑所有交易成本、竞买竞卖利差、管理时间和交易时间、努力程度、策略风险等。简而言之，回归检验所隐含的可预测性在现实世界中不可能产生风险调整后的超额利润，处理回归方程中那些统计上显著的“信息”的一种合理方法是认为它们没有经济意义。

上述回归方程为股票收益的日历效应提供了一个例子，日历效应存在很多种类。这里用传统检验方法和5%的临界值对这些效应进行简单考察。由此产生的一个问题是，发现日历效应的这些（很有可能是大量的）研究是否正确，因为这些研究考虑了一系列的日历效应（如一月效应、周内效应、周末效应），也考虑了相互替代的一些数据时期，因此为了支撑日历效应，这些研究使我们的判断发生了偏差。后面章节将讨论这一问题，这就是著名的数据调查偏

差问题。

为了在短期内（如每日交易和每天内交易）“击败市场”，人们也会使用一些技术交易规则（如表格、过滤规则、支撑和阻碍水平、神经网络、遗传算法），后面的章节将针对股票和外汇进行讨论。

风险与投资期限

大家经常使用的一个“经验规则”是，在“股票 + 债券”的投资组合中，投资者持有股票的比例应等于“100 减去他的年龄”。因此，年龄为 20 岁的投资者持有债券的比例为 20%，持有股票的比例为 80%；而年龄为 70 岁的投资者持有股票的比例为 30%，持有债券的比例为 70%。这样经验规则存在任何逻辑吗?

假设你持有诸如股票等风险资产的决策只取决于预期超额收益和投资组合收益的方差（我们将在下一章、15 和 16 章中详细讨论“均值 - 方差”方法）。股票风险短期比长期高吗?长期内的更高收益能补偿所承担的更高风险吗?我们考虑一个“基准”情形：假设连续复利收益独立同分布（这里不要求正态分布）。因此，平均收益和收益的方差都与投资期限成比例，也就是说，$E_0R_{0\to k}=k\mu$，$\sigma_{0\to k}^2=k\sigma^2$，这里的 k 为期限，比方说用年表示的时间长度，μ 是年预期收益（连续复利），σ 为年标准差。当然，方差是期限的递增函数，但预期收益也是期限的递增函数，因此，每单位方差的收益（或超额收益）对所有期限而言都相同。从而在独立同分布的假设下，我们可以认为无论期限长短，两个投资者（具有相同的风险规避程度）在风险资产上的投资比例相同（规范分析参见第 15 和 16 章）。

注意，当 $\mu>0$ 且收益服从独立同分布的标准正态分布（*niid*）时，如下结论是正确的：随着期限的增加，投资者在股市遭受损失的概率（即 $R_n<0$ 的概率，其中 R_n 是 n 年中的收益）会减小。对 *niid* 收益和 n 年期限而言，这一“不足概率”（shortfall probability）由标准正态分布变量 z_n 小于 $-n\mu/\sqrt{n}\sigma=-\sqrt{n}\mu/\sigma$ 的概率来给出，随着 n 的增加，概率值会减小。然而，即使不足概率变得非常小，比方说 10 或 20 年之后，这肯定也不是你在股票市场进行大量投资的一个充分理由，因为你的期末财富仍然只有很少增加或几乎没有增加，持有无风险的政府指数连动债券可能会更好一些，你至少确切地知道能以较高的概率“收回资金”，这个简单考虑不大可能主导你投资组合的选择。然而，这并不是说均值和方差是你投资决策所必需考的唯一因素。不同目标（如潜在损失“造成的伤害”比同等收益带来的益处要大得多，最小化“非常糟糕的货币结果”的概率）会给出不同的投资选择——正如我们在本书后面部分中要看到的那样。

与不足概率相关联（但不同）的概念是（百分比）在险价值（Var）。这是你在一个给定时期上（如 1 年）预期损失的最大值，它在 5% 的时间中（如 20 年中的 1 年）会发生。如果年平均股票收益为 μ（%），标准差为 σ（%），收益服从 *niid*，那么百分比 Var 为（$\mu-1.65\sigma$），这里的 1.65 为正态分布左尾 5% 的切断点。例如，当每年的 $\mu=6\%$，$\sigma=20\%$ 时，你预期 20 年期间 1 年内的损失最大值为 27%（在 20 年期间的 19 年中，你的预期损失小于这个值）。当然，你在任何一个年份中损失的实际数目是未知的——只有临死的人才能作出这种概率声明，只有上帝才是无所不知的。

长期收益

我们在上面讨论了 *iid* 收益的“基准”情形及其对投资组合选择的含义。什么能证明收益是 *iid* 的? 或者，长期股票收益是均值反转的吗，即更高平均收益之后是更低的平均收益? Fama 和 French（1988b）、Poterba 和 Summers（1988）发现长期（即超过两年）股票收益存在均值反转的证据。Fama 和 French 对自回归系数进行了估计，$t-k$ 到 t 区间上的收益 $R_{t-k,t}$ 与 $R_{t,t+k}$ 相关：

$$R_{t,t+k}=\alpha_k+\beta_kR_{t-k,t}+\varepsilon_{t+k} \tag{6}$$

Fama 和 French 使用的是美国总的股票指数的月度数据，考虑的收益期为 $k=1$ 到 10 年，长期数据集合覆盖了 20 世纪的绝大部分数据。他们几乎没有发现任何可预测性，除了 $k=2$ 和 $k=7$ 年（$\beta<0$）的持有期外。当 $\beta=-0.5$ 时，在 $k=5$ 处存在一个峰值，这意味着，平均来看，5 年 10% 的负收益之后的 5 年是 5% 的正收益。3～5 年期回归方程中的 R^2 约为 0.35。这一均值反转结论与“异像文献”所得出的结论一致，“买低卖高”的价值增长策略能获得收益（见第 18 章）。

Poterba 和 Summers（1988）用不同持有期收益的方差来对均值反转进行考察。如果股票收益是随机的 *iid* 变量，那么持有期收益的方差随着持有期长度的增加而增加。为了看清这一点，假设预期（对数）收益为常数，$E_t h_{t+1} \equiv E_t p_{t+1} - p_t = \mu$。在理性预期假设下，这意味着（对数）股票价格遵循随机游走模型，k 期中的收益为：

$$h_{t,t+k} = (p_{t+k} - p_k) = k\mu + (\varepsilon_{t+1} + \varepsilon_{t+2} + \cdots + \varepsilon_{t+k}) \tag{7}$$

在理性预期假设下，预测误差 ε_t 为零均值的 *iid* 序列，因此

$$\mathrm{E}_t h_{t,t+k} = k\mu \quad 且 \quad \mathrm{Var}(h_{t,t+k}) = k\sigma^2 \tag{8}$$

一种替代或一种更为直接的表示是，如果对数收益是 *iid* 序列，那么

$$\mathrm{Var}(h_{t,t+k}) = \mathrm{Var}(h_{t+1} + h_{t+2} + \cdots + h_{t+k}) = k\,\mathrm{Var}(h_{t+1}) \tag{9}$$

方差比统计量的定义是

$$\mathrm{VR}_k = \left(\frac{1}{k}\right)\frac{\mathrm{Var}(h_{t,t+k})}{\mathrm{Var}(h_{t+1})} \approx 1 + \frac{2}{k}\sum_{j=1}^{k-1}(k-j)\,\rho_j \tag{10}$$

如果收益为 *iid* 序列，那么该统计量的值为 1，在均值反转情形下（即负的自相关系数 ρ_j），该值小于 1。对 *iid* 收益而言，“调整”后的夏普比为

$$SR_{t,t+k}/\sqrt{k} = \mathrm{E}_t h_{t,t+k}/stdv(h_{t,t+k}) = \mu/\sigma \tag{11}$$

对所有期限 k 而言，这个夏普比都是常数。如果股票收益是均值反转的，那么长期比短期更安全，因为长期收益的方差上升速度小于 k。Famma－French 回归、VR 统计量和夏普比是同一现象，即均值反转的三种不同度量方式。事实上，小且自相关系数为负的收益序列就能导致长期均值反转，可以证明，VR_k（见(10)）和 β_k（见(6)）可以写成这些负自相关系数和的形式（参见 Cochrane（2001），p 1373）。

使用美国的长期数据，Poterba 和 Summers（1988）发现，对滞后期为 1 年的情形而言，方差比大于 1，对滞后期大于 1 年的情形而言，方差比小于 1，这意味着收益是均值反转的（对 $8>k>1$ 而言）。当使用许多替代性的总量股票价格指数时，尽管在对持久收益或暂时收益进行考察时检验功效较低，这一结论在一般情形中也成立。Fama 和 French（1988b）已经证明，使用 VR_k 和 β_k 得出的结论不很稳健（Cochrane（2001））。例如，对某些时期（如美国 1947—1996 年期间，使用总量股票价格指数）而言，3～5 年的 β_k 系数为正；但随着 k 的增加，“调整后的”夏普比为常数（意味着 *iid* 收益），而 VR_k 通常是一致地小于 1（意味着均值反转）。然而，对个体股票收益而言，均值反转的证据要稍微强一些。因此，对总的股票指数而言，使用不同统计量对均值反转进行测度在小样本中会得到不同的推断，尽管在长期情形中确实存在均值反转相当弱的证据。

在收益 *niid* 的零假设下，对非重叠（对数）收益而言，$\theta(k) = \mathrm{VR}_k - 1$（Lo 和 MacKinlay（1988））的极限分布为：

$$\sqrt{T}\theta(k) \sim N(0,\ 2k)$$

对重叠（对数）收益而言，

$$\sqrt{T}\theta(k) \sim N\left(0,\ \frac{2(2k-1)(k-1)}{3k}\right)$$

T 为观察值的数目，$n=k/12$，n 为年数，k 为计算 VR_k 的月份数目。然而，这些统计量在小样本中不是有效的，即使收益（即价格变化）为 *iid* 序列，VR_k 也会偏离 1。使用由历史月度收益数据校准后的 *niid* 序列的 1000 次 MCS，可以证明：当我们使用 900 个月度数据（75 年）时

（表3，Panel A 的下面部分），10 年期（重叠月度数据）VR_k 的中值为 0.810，而不是 1。因此，对 10 年期而言，$VR_k < 1$ 的点估计并不必然意味着均值反转存在。

表 3　　股票收益（不包括红利）的方差比

国家	1 年	3 年	5 年	10 年
Panel A：月度数据（1921.1～1996.12）				
美国	1.0	0.994	0.990	0.828
英国	1.0	1.008	0.964	0.817
全球	1.0	1.211	1.309	1.238
渐近检验统计量 5%，单尾	–	0.712	0.571	0.314
MCS（正态分布）				
（a）VR 中值	–	0.960	0.916	0.810
（b）第 5 个十分位数	–	0.731	0.598	0.398
Panel B：月度数据（1915.1～2004.4）（道琼斯工业平均指数）				
美国	1.0	1.198	0.886	0.549
MCS（正态分布）				
（a）VR 中值	–	1.178	0.860	0.499
（b）第 5 个十分位数	–	0.690	0.432	0.168

资料来源：Panel B，作者计算；Panel A 取自于 Jorion（2003）。

Jorion（2003）使用 30 个国家 1921—1996 年期间总股票市场指数（Goetzmann 和 Jorion（1999））发现，就所研究的 30 个国家而言，在 VR_k 的 MCS 分布（左尾 5% 的显著性水平）基础上，没有任何证据证明 1～10 年期上的实际收益是均值反转的。在表 3 中，随着期限的延长，英国和美国的方差比统计量倾向于减小，但减小的速度不快。无论计算的股票收益是否包含红利，结论都不会发生变化。就曾经中断过的市场而言（如俄罗斯、德国、日本），倾向于均值厌恶（mean aversion）（即 $VR_k > 1$），这意味着全球指数也呈现出均值厌恶的特征（表 3，Panel A，（a）和（b））。渐近检验统计量和正态分布假设下由 MCS 给出的 5% 显著性水平下的单尾临界值表明，方差比统计量会拒绝均值反转的零假设。

就美国而言，我们使用道琼斯指数（工业平均指数，1915 年 1 月～2004 年 4 月期间非重叠的月度数据），将最新数据加入到 Jorion 的数据集合中去，在很宽的范围内发现了类似的结果（表 3，Panel B）。就 5 年期和 10 年期而言，与 Jorion 的数据相比，均值反转的证据要稍微明显一些。在无均值反转效应（和正态分布）的零假设下，VR_k 统计量的中值再一次向下偏斜，但是，由实际数据计算的 VR_k 统计量值比由 MCS 中 5% 的单边临界值要大，这证实了如下结论：我们能拒绝长期中均值反转效应存在的零假设。

根据波动率的大小，可以认为，在绝大多数时期中，许多国家的股市都比美国股市的风险高。就 5 年期而言，Jorion 的研究所发现的一些波动率值（对数价格变化）为美国 59.7%，奥地利 82%，芬兰 86%，意大利 78%，日本 89%，德国 94%，葡萄牙 155%，菲律宾 212%，巴西 228%，智利 250%，哥伦比亚 100%，墨西哥 187%。

通过使用美国实际收益的真实经验分布（而不是假设收益服从正态分布），Jorion 发现，当使用价格变化（即资本收益）来计算股票收益时，对 1 年、5 年、10 年期而言，美国股票遭受损失的概率只有轻微的下降，分别为 36.6%、34.3% 和 33.7%（表 4）。当我们将 3.8% 的红利收益包含进来后，对总收益（即资本收益加红利收益）重新进行计算，正如我们所预期的那样，对 1 年、5 年、10 年期而言，不足概率有所下降，分别为 30.8%、20.7%

和 15.5%。

这一组数值中哪些能更好地代表长期风险呢？包括红利的一些数值取自于没有关闭过的股票市场，因此收益率序列存在大于 150 个基点的幸存偏差。不足概率并没有考虑投资者始终可以投资国库券或债券的事实，因此使用超额收益来度量不足概率可能要更好一些。假设国库券的实际收益为正但很低，比方说每年 1%，这会抵消一部分红利收益，增加超额收益的不足概率（相对于股票总收益的超额收益而言）。

表 4　　长期风险与收益：1920—1996

	损失概率			Var（5%，左尾）		
	1 年	5 年	10 年	1 年	5 年	10 年
美国（价格变化）	36.6	34.3	33.7	-27.8	-45.5	-51.2
美国（总收益）	30.8	20.7	15.5	-24.5	-33.7	-22.3
英国（价格变化）	40.3	32.5	45.2	-24.5	-54.8	-50.8
英国（总收益）	30.1	22.1	30.8	-24.9	-47.8	-45.1
中值（价格变化）——30 个国家	48.2	46.8	48.2	-31.0	-60.3	-65.4
中值（总收益）——15 个国家	36.1	26.9	19.9	-24.7	-39.9	-34.8
全球指数（价格变化）	37.8	35.4	35.2	-20.8	-40.4	-41.1
全球指数（总收益）	30.2	18.2	12.0	-16.7	-19.8	-11.2

注：

（1）全球指数用 GDP 进行加权。

（2）总收益为价格变化的百分比加上红利收益的百分比。

（3）统计量的值根据收益的经验分布来进行计算。

（4）使用总收益的中值变化的覆盖范围要窄一些，因为只有 15 个国家的红利数据是可以获得的。

我们也必须考虑交易成本和管理费用。对指数（追踪型）基金而言，（目前的）管理费用约为 50 ~ 100 个基点，你所选择基金的表现当然会低于指数（根据 Frino 和 Gallagher（2001），就标准普尔 500 指数基金而言，平均成本为 40 ~ 50 个基点，因此，国外指数基金的平均成本可能会更高一些）。就管理型（共同）基金而言，交易成本可能在 300 ~ 500 个基点的范围内，平均而言，管理型基金的表现会低于指数基金（就共同基金的绩效而言，参见第 9 章）。因此，这些交易成本将 3.8% 的红利收益部分或全部抵消掉，考虑交易成本之后的不足概率更接近于只使用资本收益时的值。

我们也必须注意的是，美国 5 年期和 10 年期的最大损失分别为 -56.4%（1934 年的 9 月末）和 -60.7%（1982 年的 8 月末）——这些是如此长时期中的“最大值”，可以说这也是全世界表现最好的市场。就使用资本收益（总收益）的美国股票而言，5% 的左尾 Var 值（即最大预期损失，在 5% 的时间中发生——见表 4 和第 28 章）为：1 年期，27.8%（24.5%）；5 年期，45.5%（33.7%）；10 年期，51.2%（22.3%）（Jorion 2003）。当我们使用道琼指数（1915 年 1 月到 2004 年 4 月）将最新数据包含进去后，Var 值与 Jorion 发现的值相似，最近的衰退期对股票的最大可能损失并没有造成大的影响。将不包括美国的 29 个国家合在一起进行考察，发现风险要更高一些（表 4 中“中值数字”可以反映这一点）。

如果我们接受长期内风险不会削弱的观点，那么这意味着，在预期收益和风险之间进行权衡的年轻投资者持有股票的比例不会比年龄大的投资者高。（当然，为什么年轻的投资者比年老的投资者持有股票的比例更高一些，可能存在其他方面的原因——参见第 16 章。）

然而，这可能是一种安慰，但安慰的程度有限。地理位置不同的 30 个国家最大损失发生的时间并不完全同步。一些国家的最大损失分别发生在 20 世纪 30 年代、二战时期、1974 年衰

退时期中，对一些发展中国家而言，最大损失发生在20世纪80年代的高通货膨胀时期及最近的2000—2003年衰退时期中。因此，全球收益指数（只用资本收益）在一年期中的波动率为16.1%（美国为20.89%）。然而，就5年期和10年期而言，全球指数波动率分别为52.87%（美国为59.74%）和80.73%（美国为81.79%），这与美国的数字之间没有多大的差异。就5年期和10年期而言，全球指数的损失概率分别为35.4%（美国为34.3%）和35.2%（美国为33.7%），这与美国的数字基本相同，而就5年期和10年期而言，全球指数的Var分别为40.4%和41.1%，这比美国的数据要稍微低一些（表4）。因此，全球分散化对美国投资者的福利微乎其微，对美国以外国家的投资者的福利可能要更大一些。

尽管支持均值反转的证据存在一些不确定性，这不必然排除可预测性。VR_k 和 β_k 都是对可预测性的单变量检验，即使 $R_{t,t+k}$在任何 $R_{t-k,t}$给定的基础上是无法预测的，多变量回归方程中的其他变量（如红利价格比、利率）也会对它产生影响。Cochrane（2001）非常聪明地证明了一个合适的向量自回归（Var）模型是如何产生低的单变量均值反转的，在该Var模型中，预期收益 h_{t+1}移动的速度很缓慢，预期收益由缓慢移动的价格红利比来决定。Var模型导致了如下形式的单变量移动平均表示：

$$\begin{aligned} h_t &= \frac{(1-\gamma L)}{1-bL} v_t \\ &= v_t - (\gamma - b)\ v_{t-1} - b\ (\gamma - b)\ v_{t-2} - b^2\ (\gamma - b)\ v_{t-3} + \cdots \end{aligned} \tag{12}$$

如果 $\gamma > b$，那么正的冲击 v_t 引发了一连串的负收益，这就导致了单变量收益序列中的均值反转。但是，这最终是由影响预期收益的移动缓慢的变量所导致的，因此，这并没有违背EMH。Cochrane的Var模型为：

$$h_{t+1} = x_t + \varepsilon_{h,t+1} \tag{13a}$$

$$x_{t+1} = bx_t + \varepsilon_{x,t+1} \tag{13b}$$

$$\Delta d_{t+1} = \varepsilon_{d,t+1} \tag{13c}$$

其中，x_t 是缓慢移动（$v > 0.95$）预测变量（如红利价格比的对数），h_{t+1}为收益，红利增长是不可预测的。Cochrane认为，当 x_t 决定 $E_t h_{t+1}$且红利增长冲击 ε_{dt}与预期收益冲击 ε_{ht}之间的相关性为零时，（12）中（$1-\gamma$）/（$1-b$）$=0.7$，这大概是数据中所发现的长期均值反转。因此，“强的”多变量可预测性与小的单变量均值反转不会相互排斥。

检验功效

如下情形始终可能发生：在样本数据中，即使正确模型中的收益是可预测的，一些特殊检验统计量也无法拒绝随机变量的零假设。“检验功效”就是拒绝错误的零假设的概率。Poterba和Summers（1988）给出了如何评估“功效”的一个漂亮而又简单的教学例子。为了对统计量 ρ_k、VR_k 和 β_k 的功效性质进行评估，Poterba－Summers建立了一个股票收益相互关联的正确模型。MCS用来生成拟合的收益数据，并计算随机检验统计量的值。如果这些“统计量”是有功效的，那么应在绝大多数情形中拒绝随机零假设。正确模型是股票收益服从ARMA（1，1）过程。假设实际收益的对数 p_t 包含持久部分 p_t^* 和暂时部分 u_t，持久部分遵循随机游走：

$$p_t = p_t^* + u_t \tag{14a}$$

$$p_t^* = p_{t-1}^* + \varepsilon_t \tag{14b}$$

暂时部分 u_t 具有持续性：

$$u_t = \rho u_{t-1} + v_t \tag{14c}$$

其中，u_t、ε_t 和 v_t 不能同期相关。Poterba－Summers设定 $\rho = 0.98$，这意味着暂时价格部分的创新可以持续2.9年的一半时间（如果基准时间区间为一月）。由上述三个方程，我们可以得到

$$\Delta p_t = \rho \Delta p_{t-1} + \ [\ (\varepsilon_t - \rho \varepsilon_{t-1})\ +\ (v_t - v_{t-1})\ ] \tag{15}$$

因此，模型意味着对数收益服从ARMA（1，1）过程。通过对误差进行随机抽取，它们

就生成了 p_t 的拟合数据，Δp_t 方差的相对份额由 σ_ε^2 和 σ_v^2 的相对大小来确定。接下来，他们根据所生成的720个观察值来计算统计量 ρ_k、VR_k 和 β_k 的值（即与收益的历史数据有相同的"长度"）。他们重复上述操作25 000次后就得到了感兴趣的三个统计量的频率分布。他们发现在区分随机游走模型（即错误的"零假设"）和上述"正确模型"（包括很大的持续性部分和暂时性部分）方面，三个统计量的功效都很弱。

自回归移动平均（ARMA）

如果弱式有效不成立，那么实际收益 R_{t+1} 不仅仅取决于历史收益，还取决于历史预测误差（见上面的(15)），正如一般的ARMA（p，q）模型：

$$R_{t+1}=k+\gamma(L)R_t+\theta(L)\varepsilon_{t+1} \tag{16}$$

其中，$\gamma(L)$ 和 $\theta(L)$ 是滞后多项式算子，例如，$\gamma(L)=1+\gamma_1L+\gamma_2L^2+\cdots+\gamma_pL^p$，$L^nR_t=R_{t-n}$。在EMH下，我们期望 $\gamma(L)$ 和 $\theta(L)$ 中所有参数都为零。基于ARMA模型的回归经常被用来对EMH假设下的信息效率进行检验。事实上，Poterba和Summers试图用ARMA（1，1）模型来拟合所生成的股票收益数据，当然，根据构造，它应该能拟合这些数据。然而，在他们的估计方程中，发现 $\gamma=0.98$，$\theta_1=1$，因为 ρ 和 θ 应非常"靠近"，所以估计程序通常是无法"区分的"（视为一样），也就无法从统计上成功地估计取值不同的 ρ 和 θ。但Poterba和Summers确实得到了这些估计值，最多10%的回归参数值接近于（已知的）正确值。这是估计模型在有限样本中不能模拟正确模型的另外一个例子。

Cecchctti、Lam和Mark（1990）从最后一点开始着手，他们对Poterba和Summers（1988）、Fama和French（1988b）表示怀疑，认为股票价值存在均值反转应用噪声交易者的存在来进行解释。他们注意到，从本质来看，收益的序列相关性并不意味着效率假设被违背。Cecchctti等继续证明，均值反转的经验证据与由时变预期收益的均衡模型所生成的数据是一致的。他们将以消费为基础的资本资产定价（CCAPM）模型（见第13章）进行特殊的参数化处理之后作为代表性均衡模型，并使用蒙特卡罗方法来生成拟合数据集合。接下来，他们对拟合数据集合使用Poterba－Summers方差比检验和Fama－French长期收益回归方法。他们发现，在用历史收益计算的股票价格中，均值反转的度量值始终落在由CCAPM模型所隐含的蒙特卡罗分布中值60%的置信区间内。与所有的蒙特卡罗研究类似，结果是均衡模型所选择的特殊参数的结果。Cecchctti等注意到，在Lucas（1978）的均衡模型中，消费等于产出，也等于红利，他们用蒙特卡罗方法对三个相互替代的"基础变量"进行了考察。以红利为例，蒙特卡罗模拟假设：

$$\ln D_t=\ln D_{t-1}+(\alpha_0+\alpha_1S_{t-1})+\varepsilon_t$$

S_{t-1} 为马尔可夫转换变量（参见第7节和Hamilton（1989）），转移概率为

$$Pr(S_t=1\mid S_{t-1}=1)=p$$

$$Pr(S_t=0\mid S_{t-1}=1)=1-p$$

$$Pr(S_t=0\mid S_{t-1}=0)=q$$

$$Pr(S_t=1\mid S_{t-1}=0)=1-q$$

因为 α_1 必须限制为负数，所以 $S_t=0$ 为"高增长"状态，$E_t\Delta\ln D_{t+1}=\alpha_0$；$S_t=1$ 为低增长状态，$E_t\Delta\ln D_{t+1}=\alpha_0+\alpha_1$（$\alpha_1<0$）。因此，$\ln D_t$ 遵循随机游走，随机飘移为（$\alpha_0+\alpha_1S_{t-1}$）。应用极大似然估计法来对红利过程的参数进行估计，然后就可以用这些参数来生成红利的拟合数据。

CCAPM模型（用红利代替消费，参见13章）的欧拉方程为：

$$P_tU'_t(D_t)=\delta\mathrm{E}_t[U'(D_{t+1})(P_{t+1}+D_{t+1})]$$

偏好由幂效用函数 $U(D)=(1+\gamma)^{-1}D^{1-\gamma}$ 给出，常相对风险规避系数（CRRA）满足 $-\infty<\gamma\leqslant0$。欧拉方程可以写为：

$$P_tD_t^{\gamma} = \delta \mathrm{E}_tP_{t+1}D_{t+1}^{\gamma} + \delta \mathrm{E}_tD_{t+1}^{\gamma+1}$$

向前重复迭代后就得到解

$$P_t = D_t^{-\gamma} \sum_{i=1}^{\infty} \delta^i \mathrm{E}_tD_{t+i}^{1+\gamma}$$

我们将问题稍微简化一下，使用上述方程（δ 和 CRRAγ 都取代表值）得到的红利拟合数据给出的是所生成的价格数据序列，它满足 CCAPM 这一一般均衡模型。所生成的收益数据 $R_{t+1} = [(P_{t+1} + D_{t+1})/D_t] - 1$ 接下来可以用于计算 Poterba - Summers 方差比统计量的蒙特卡罗分布，并进行 Fama - French 的长期收益回归估计。

从本质来看，Cecchetti 等的结果说明，使用美国可获得的 116 个年度历史观察值，我们对基于 5 年期收益所进行的回归方程不会有多大的信心，因为只有 23 个非重叠数据。因此，历史数据太短以至于我们无法在均衡模型和“噪声交易”模型之间进行清晰的选择。Cecchetti 等人的结果在解释弱式有效方面确实要谨慎得多（只用滞后收益数据），这为噪声交易者的存在提供了信号。

4.4 多变量检验

Fama - French 和 Poterba - Summers 的结果都是单变量检验。然而，人们发现，除历史收益之外的许多变量都有助于预测当前收益。早期研究，如 Keim 和 Stamburgh（1986），使用美国 1930—1978 期间股票（相对于国库券利率）的超额收益数据发现，对许多投资组合而言（以规模为基础），如下变量（有些随意）通常是显著的：（*a*）低等级公司债券之间的收益差异和月度国库券收益；（*b*）在过去 4 ~ 5 年中，上期（实际）标准普尔指数相对于平均值的偏离；（*c*）只包含“小盘股”的股票价格指数。

他们也发现了“纯粹”的一月效应，因为针对一月的（*a*）~（*c*）的影响与针对其他月份的结果不一样。（他们还发现，上述变量会对诸如政府长期债券、高等级公司债券等资产的月度收益率产生影响。）然而，应该注意的是，就股票的月度收益数据而言，回归方程只能解释实际超额收益的 0.6% ~2%。

Fama 和 French（1988b）对（名义和实际）收益与红利收益比 D/P 之间的关系进行了考察。

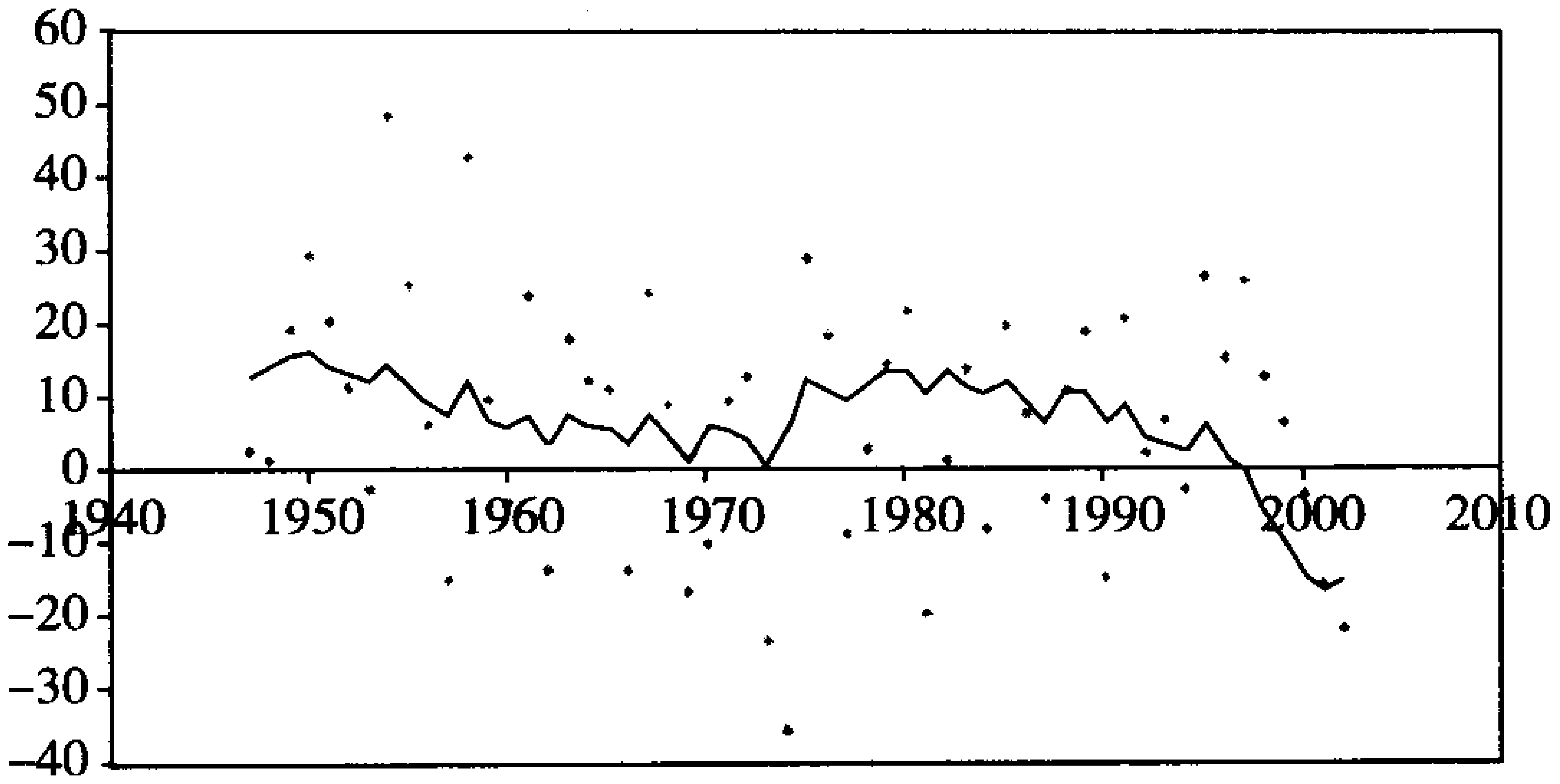
$$R_{t,t+k} = a + b\ (\mathrm{D/P})_t + \varepsilon_{t+k} \qquad (17)$$

方程用 1 ~4 年期月度或季度 NYSE 指数收益来进行回归。通过在各个子时期上进行回归，他们也对方程的稳健性进行了检验。就月度和季度数据而言，红利收益比在统计上经常是显著的（且 $b>0$），但只能解释实际收益 5% 的方差。对更长的时期而言，解释功效有所增加。例如，就 1941—1986 年期间的名义收益而言，1、2、3、4 年期的解释力分别为 12%、17%、29%、49%。人们也发现更长收益期的回归可以用来进行“样本区间外”预测。

可预测性与择时交易

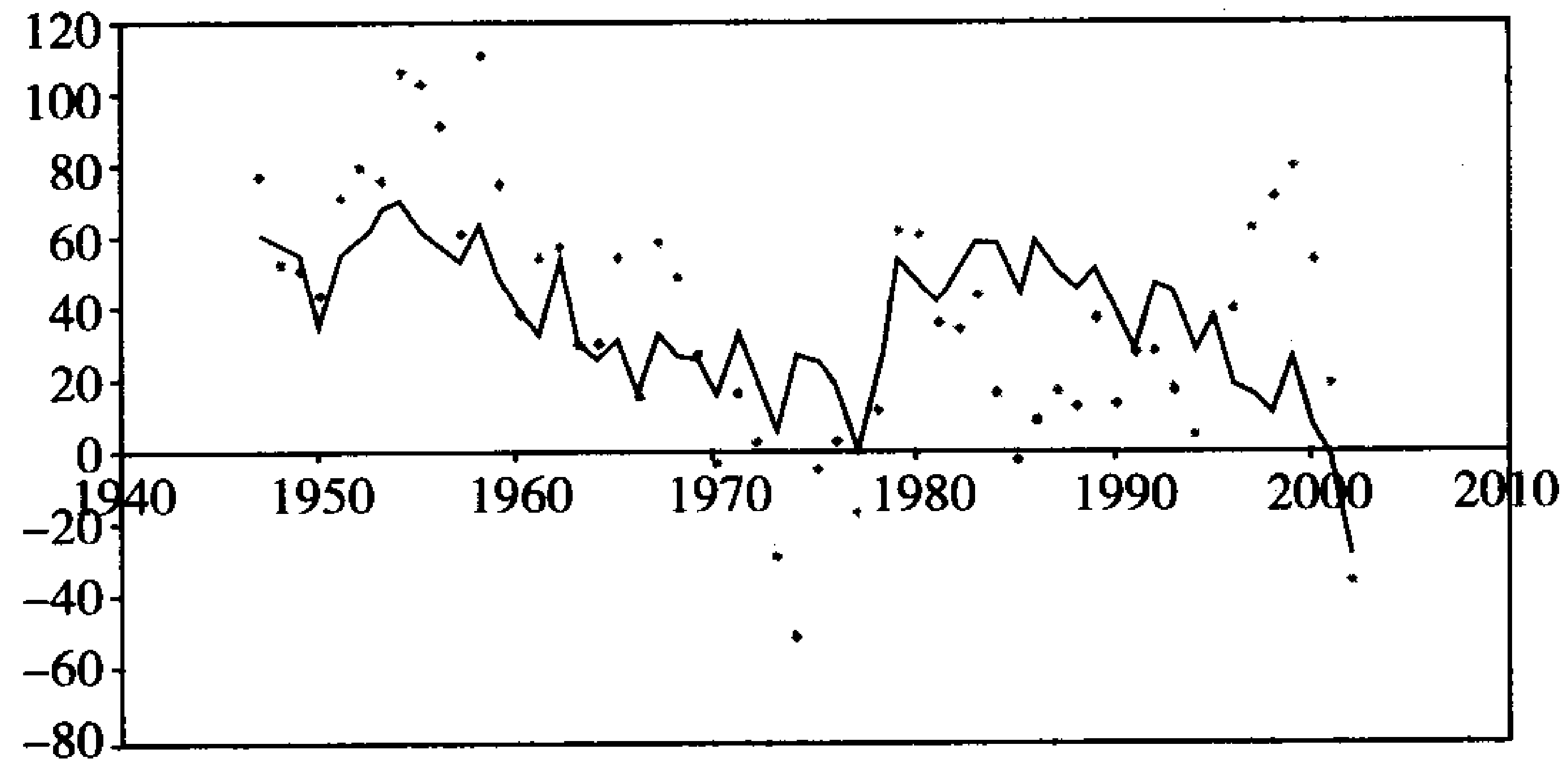
最近，Cochrane（2001）使用 Shiller 的超额收益数据——美国 1947—1996 年的一年期数据——发现 $b \approx 5$（标准差为 2），$R^2 = 0.15$；而对 5 年期数据而言，$b \approx 33$（标准差为 5.8），$R^2 = 0.60$。我们已经使用过 Shiller 关于美国 1947—2002 年的数据，实际收益、预测（样本内）的年度收益、5 年期的超额收益如图 6 和图 7 所示，数据结束日期为 2002 年。

显然，一年期收益具有很强的波动性，价格红利比只能解释收益变化很小的一个部分。粗略地看，5 年期收益似乎能更好地拟合数据。价格红利比（图 8）是一个运动缓慢的（持久的）变量，每 20 年与均值相交一次，在过去的 10 年中（到 2002 年为止），该变量上升到史无前例的 70 左右。



美国:1 年期收益:1947—2002(实际,拟合)

图 6　1 年期超额收益



美国:5 年期收益:1947—2002(实际,拟合)

图 7　5 年期超额收益

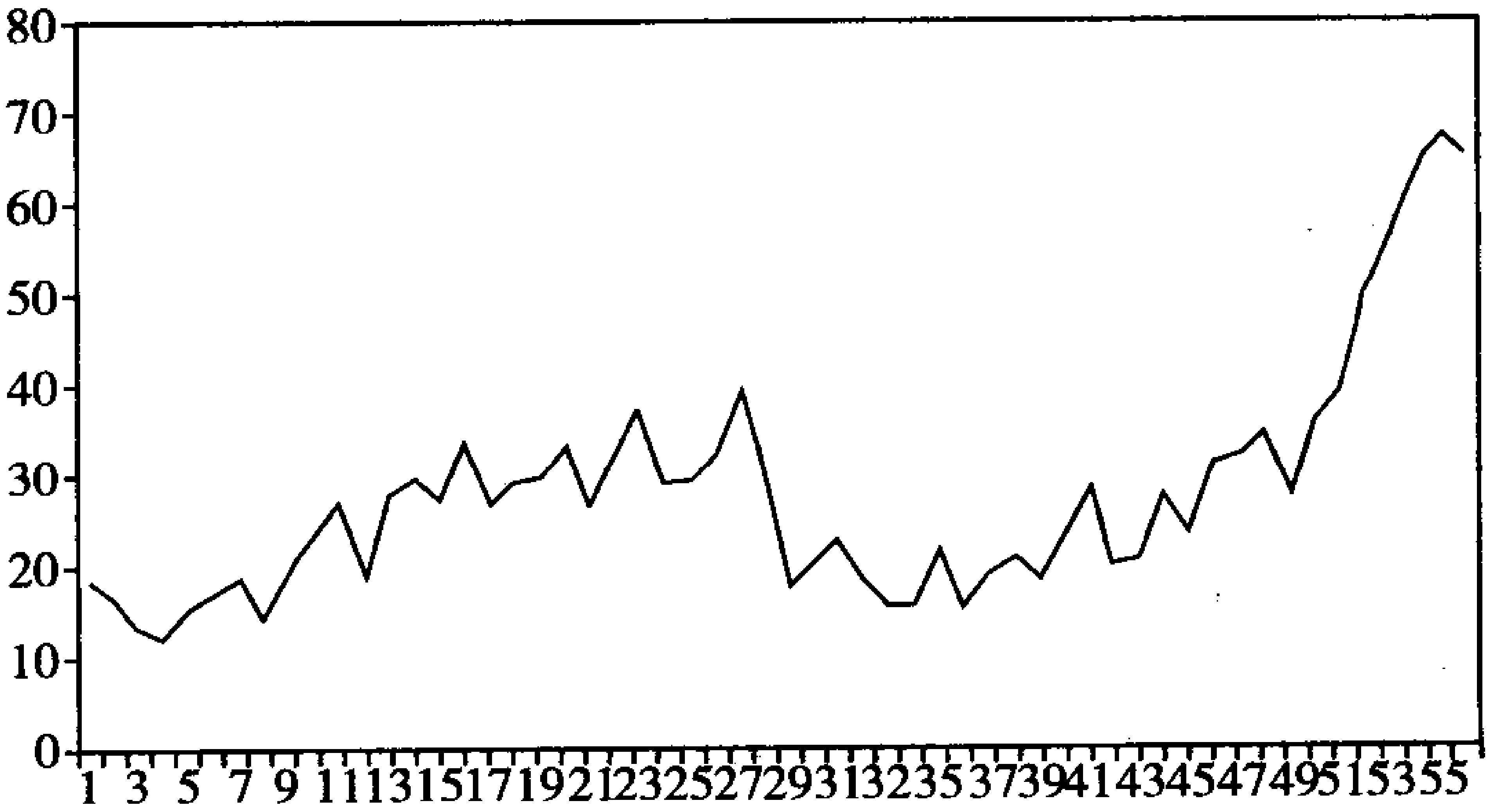
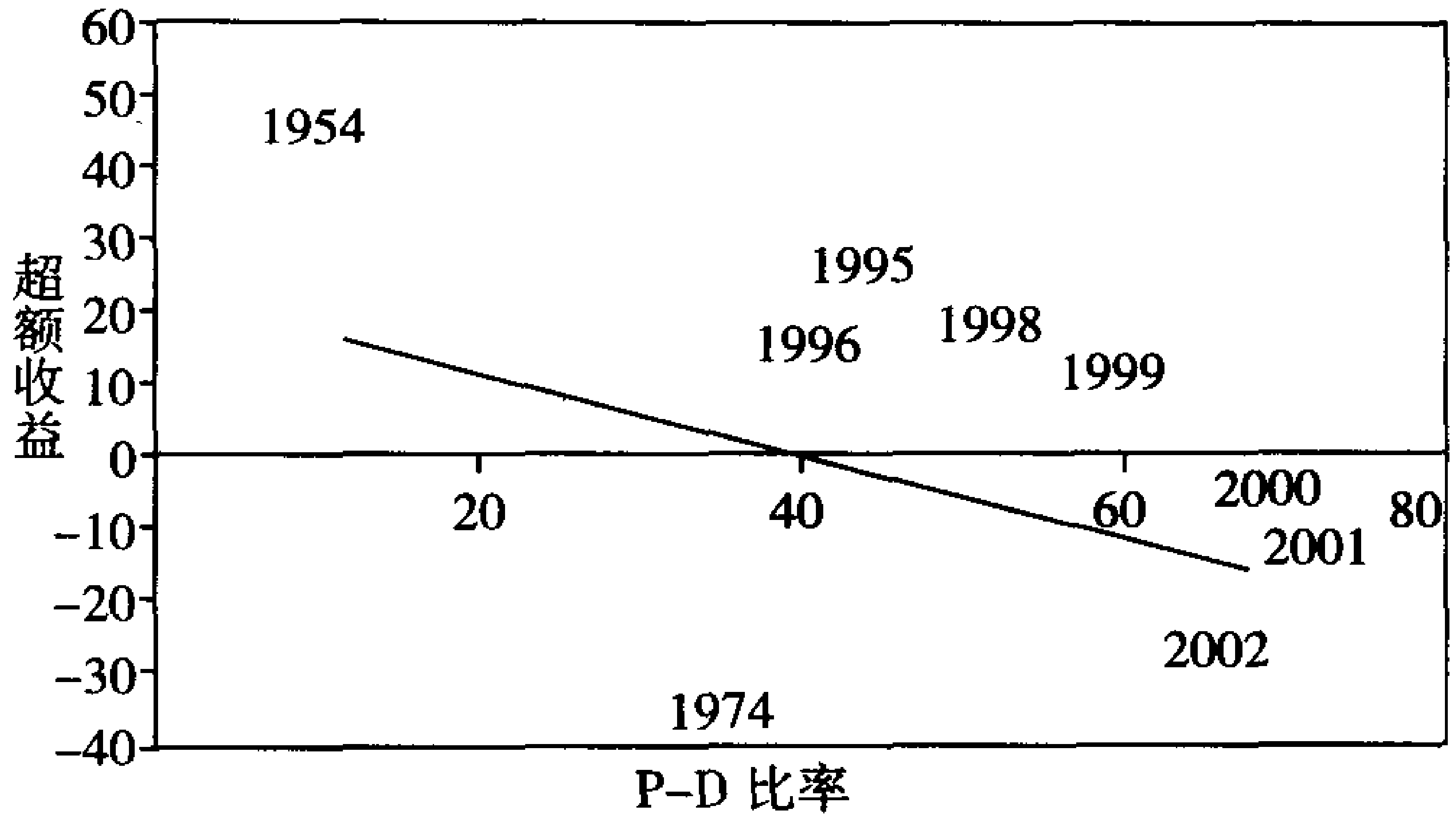


图 8　价格红利比率：美国（1947—2002）



一年期超额收益和价格红利比：美国年度数据（1947—2002）

回到一年期收益上（图9），我们可以看到，方程预测 1998 年和 1999 年的收益为负（样本内预测），但这些年份上出现的是价格大幅上涨（部分原因可能是电信业的繁荣）。如果你用这个模型的计算结果来选择进入市场的时机，那么你将错过两年价格大涨的机会，也会过错 1995 年价格大涨的机会（如果你是一个职业交易者，你可能已经被解雇了）。在 2001 年和 2002 年，方程又“回到轨道上来”了。

样本外预测的表现可能比上面讨论的样本内预测更差一些。事实上，Cochrane（1997）使用直到 1996 年的数据（他所拥有的最新数据）来对这些方程进行估计，并用 AR（1）方程来对价格红利比（P/D）进行预测：

$$R_{t+1} = a + b\ (P/D)_t + \varepsilon_{t+1} \qquad b<0$$

$$(P/D)_{t+1} = \mu + \rho\ (P/D)_t + v_{t+1}$$

使用第二个方程来对 $(P/D)_{1997} = \mu + \rho\ (P/D)_{1996}$ 和 $\hat{R}_{1998} = a + b\ (P/D)_{1997}$ 等进行预测，Cochrane 发现，方程预测 1997 年的超额收益为 -8%，10 年之后的预测值仍为每年 -5%。显然，图 9 表明，这些预测值在 1998 年和 1999 年的误差非常大，这两年出现的是非常高的收益，但 2000—2002 年期间的“样本外”预测却“回到了原来的轨道上来”。当然，如果假设价格红利比在 20 世纪 90 年代高于其历史平均值，价格红利比确实如数据集合所表明的那样是缓慢运动的（即自相关系数 $\rho > 0.95$），那么收益方程会预测未来几年的收益为负。但图 8 和图 9 说明了对收益可预测性进行评估过程中的一些问题。记住，ρ 是有偏的，由测度误差来度量，因此，价格红利比的持续性小于给定的点估计值。如果价格红利比根本不具有持续性（显然这不是一个正确的结论），那么它会迅速返回到均值上来，所预测的任意负收益只能持续一个时期。

可能的情形是，价格红利比的线性 AR 模型不正确，而非线性模型是合适的，如果当前状态与均值有很长的距离，那么价格红利比以更快的速度回到均值状态。（我们假设向长期均值水平运动的绝大部分是由价格变化而不是红利变化所造成的——因为人们已经发现价格红利比对未来红利变化有一点或几乎没有可预测性，Cochrane（2001）。）我们在这里设定了价格红利比的一个非线性动态反应过程，下一节将对此进行讨论。这意味着 2002 年以后几个年份中的收益为负。

另一方面，也许从现在开始，价格红利比的长期平均值比历史平均值高（如由“新经济”引起的生产率增加方面的原因，未来利润在一段时间内会有所增加，因此相对于红利而言，价格也会有所增加）。这也意味着，未来收益不会快速下降，因为价格红利比在到达新的更高均值之前，下降的距离要更短一些。

当我们考察收益的回归方程时（图 9）很有可能出现的情形是，R_{t+1} 与 $(P/D)_t$ 之间是一种非线性关系，如果“直线”在右边部分（即 2000—2002 年的数据点上）非常扁平，那么拟

合效果会更好一些。如果是这种情形，那么新方程只能预测未来价格变化一个非常小的下降。为了强调这些结果的脆弱性，值得注意的是，Ang 和 Bekaert（2001）在对全世界的股票市场进行研究后发现，当20世纪90年代的数据包含在样本中时，红利价格比对股票收益有一点或几乎没有任何预测功效，短期利率成为股票收益最为稳健的预测指标（也可参见 Goyal 和 Welch（1999））。总而言之，这些证据会让你对如下事实产生警惕：“择时交易”是一个很冒险的职业。

同样应该保持警惕的是，表示长期内“可预测性更强”的那些统计量会在两个方面形成误导。第一，“b”是有偏的（Stamburgh（1999））。第二，在预测变量（D/P）具有持续性的那些地方，从逻辑角度考虑，如果短期内收益的可预测性很弱，回归系数“小”，那么长期收益回归方程中“b”的值会增加，R^2 也会更大。为了看清前面一个结论，考虑

$$R_{t+1} = bx_t + \varepsilon_{t+1} \tag{18a}$$

$$x_t = \rho x_{t-1} + v_t \tag{18b}$$

其中，x_t 的持续性非常强（即 $\rho > 0.9$），x_t 代表预测变量（即 $x_t \equiv \ln(D/P)_t$）。例如，三期收益（连续复利）为

$$(R_{t+1} + R_{t+2} + R_{t+3}) = b(1+\rho+\rho^2)\ x_t + b(1+\rho)\ v_{t+1} + bv_{t+2} + \sum_{i=1}^{3} \varepsilon_{t+i} \tag{19}$$

因此，对 $\rho = 0.9$ 和 $b > 0$ 而言，x_t 的系数是期限的递增函数（与 R^2 一样——这里没有给出结果）。R^2 随着期限的增加而增加，这让人感到有些不牢靠，这大致说明，从长期来看，红利价格比高的那些时期之后倾向于出现的情形是，下一个五年的收益也很高。但是，在使用年度数据和五年期收益时，重叠数据问题意味着我们实际上只用了1/5的观察值，因此当我们只对一般统计量进行考察时，可预测性问题比看上去的情形还要严重。

对估计系数的稳定性而言，什么才是最重要的呢？当然，为了进行如下研究，你没有必要进行更为深入的考察：诸如红利（收益）价格比率、（没有趋势的）国库券利率、收益率差、违约收益率差（即公司收益减去国库券收益）、消费－财富比率等收益预测变量在有些时候有助于“解释”（不同时期上）收益的时间序列行为，但这些变量的系数并不特别稳定（参见 Lettau 和 Ludvigson（2001a）、Pesaran 和 Timmermann（2000））。这意味着，在解释“可预测性”时，我们必须非常小心，已有大量研究对各种变量和假设进行过考察（即数据挖掘）。显然，为了考察回归是否真正地提供了“择时交易”的优势，“样本外”预测检验在这里是很有用的。

EMH 意味着超额收益不可预测，而不是实际收益不可预测。几个研究发现了“可预测性”，但是，因为它们所使用的均衡收益模型不合理，所以我们不知道 EMH 在一般情形中是否应该被拒绝。例如，（17）中 $b > 0$ 的发现可以在 EMH 下用如下断言来进行解释：（D/P）是均衡预期收益变化的一个替代变量（有理论对此进行支持——见第12和13章）。再举一个例子，与 Keim 和 Stamburgh 的研究一样，低等级债券收益的增加反应了投资者对“风险”感知的增加，我们预期均衡收益和真实收益两者都会发生变化。可预测性在这里令人信服地与 EMH 一致，尽管这些事后解释在无一致均衡收益理论模型时显得有些无力。

4.5　协整与误差修正模型

正如下一章即将看到的那样，股票价格 P 由预期未来红利 D 的现值和贴现率（即单期未来收益）来决定。因此，红利价格比对未来红利或未来收益或两者同时具有预测功能。如果预期未来收益不为常数，那么我们就有一个理论上的理由去相信价格红利比可以预测未来收益。

假设长期（静态）均衡中的价格 $p = \ln P$ 与 $d = \ln D$ 成比例，价格红利比为常数（$= k$）。

$(p-d)$ 对 k 的偏离可能导致价格变化。这就是 ECM 的基础。我们假设价格在将来变化的速度和方向取决于“长期”红利价格比中非均衡的程度。如果“长期”（对数）价格红利比 $z=(p-d)$ 被假设为常数，那么“标准的”误差修正模型（ECM）为

$$\Delta p_t = \beta_1 (L)\ \Delta d_{t-1} + \beta_2 (L)\ \Delta p_{t-1} - \alpha\ (z-k)_{t-1} + \varepsilon_t, \qquad \alpha > 0 \qquad (20)$$

其中，$\beta_i\ (L)$ 滞后多项式算子（$i=1,\ 2$），k 是（对数）价格红利比的长期均衡值（年度数据中的 P/D 约为 20）。当价格高时（相对于“长期”红利而言），$(p-d) > k$，那么 Δp_t 为负，且价格在下一期会下降，这使 $(p-d)$ 向均衡值回移。如果 p_t 和 d_t 是非平稳的 I（1）序列，但两者具有协整关系，那么 $(p-d)_{t-1}$ 是 Δp_t 或 Δd_t Granger 原因。这就是与（20）类似的预测方程的统计基础。

MacDonald 和 Power（1995）用 1871—1987 年美国的年度数据对像（20）那样的 ECM 进行了估计。为了获得一个平稳的协整向量，他们不得不将持久比率（等于保留收益/总收益）和红利变量包含进去。他们在样本内发现了可预测性的证据，得到了合理的样本预测值（1976—1987），尽管后者很有可能由同期的 Δd_t 所造成。类似地，Clare、O'Brein、Thomas 和 Wickens（1993）在金边股本收益比率（gilt - equity yield ratio，GEYR）的基础上给出了 ECM 模型。GEYR 被市场分析师广泛用于股票预测。$\text{GEYR}_t = (C/B_t)\ /\ (D/P_t)$，其中，$C$ = 永久债券的息票支付，B_t = 债券价格，D = 红利，P_t = 股票价格（金融时报股票价格指数）。有人争论认为，英国的大机构投资者——养老基金——在短期内关心的是收入流而不是资本收益。因此，当 (D/P) 相对于 (C/B) 较低时，他们会售卖股票（购买债券），从而高 GEYR 预示下期股票价格会下降。从宽泛意义上讲，市场分析师所使用的经验规则是，如果 GEYR > 2.4，那么就售卖股票；如果 GEYR < 2，那么就购买更多的股票；如果 2 ≤ GEYR ≤ 2.4，那么就“保持”头寸不变。能刻画上述思想的一个简单模型是：

$$\Delta \ln P_t = \alpha_0 + \beta \text{GEYR}_{t-1} + \sum_{i=1}^{3} \alpha_i \Delta \text{GEYR}_{t-i} + \varepsilon_t$$

在静态模型中，长期 GEYR 为 $-\alpha_0/\beta$，Clare 等（使用季度数据）发现，GEYR 等于 2.4。GEYR 比 2.4 每增加 0.1，估计方程所预测的金融时报股票指数会下降 5%：

$$\begin{aligned}
\Delta \ln P_t &= \underset{(3.7)}{122.1} - \underset{(3.4)}{50.1}\text{GEYR}_{t-1} + \underset{(2.4)}{48.1}\Delta\text{GEYR}_{t-1} + \underset{(3.4)}{6.5}\Delta\text{GEYR}_{t-2} + \\
&\quad \underset{(2.8)}{32.6}\Delta\text{GEYR}_{t-3} + \text{虚拟变量}\ [73\ (3),\ 73\ (4),\ 75\ (1),\ 87\ (4)] \\
&\quad 1969\ (1)\ \sim 1992\ (2),\ \bar{R}^2 = 0.47
\end{aligned}$$

他们对上述方程进行迭代估计，然后对 1990（1）～1993（3）上的结果进行预测，假设投资者的交易规则是：如果预测的资本收益大于三个月国库券的利率就持有股票（否则就持有国库券）。与前面的经验规则分析师情形相比（年收益为 0.4%，标准差为 $\sigma = 39.9\%$），基于上述估计方程的这种策略所得出的事后资本收益要更高一些，为每年 12.3%，而标准差要小一些（$\sigma = 23.3\%$）。因此，该研究提供了违背弱式有效的表面证据；注意，分析中使用的是资本收益而不是总收益，且没有考虑交易成本。但这两个因素不大可能破坏他们结论的基础。

就美国和英国的情形而言，上述分析是对 Harris 和 Sanchez - Valle（2000a，b）的重复和扩展，只不过使用的是截止到 1997 年的月度数据，上述 GEYR 交易规则所产生的收益比买而持有策略（即使考虑交易成本）的收益每年高约 1%～3%，但这一结果没有对事前风险进行校正。Clare、Priestly 和 Thomas（1997）使用英国的 GEYR 来预测 FTSE100 的收益，但没有发现样本外的可预测性。因此，GEYR 能用于“择时交易”，从而击败对风险调整后的市场的证据相当脆弱。

协整和 Engle - Granger 定理被成千上万的实证研究用来从一组非平稳的 $(n\times 1)$ I（1）变量出发来构建一组协整向量（最多 $n-1$ 个），这些研究使用单方程 ECM（或 Johansen（1988）的 Var）或多变量 VECM 来考察变量 x 中一些或全部的可预测性（即 Granger 因果关系）。数据

挖掘的范围在这里很宽泛，人们对这些结果的信心部分取决于统计分析师的好坏（如合适诊断、样本外预测等），也取决于回归背后的“理论”模型。

例如，金融领域有大量研究来考察即期汇率的可预测性，预测基础要么是购买力平价偏差和远期溢价，要么是已经出现的利率平价。股票价格可预测性的许多研究都是基于股票价格和红利（或收益）之间的协整分析，这些研究导致了 VECM 模型的使用，VECM 模型中用某个方程来“解释”股票收益。其他研究是基于不同股票市场指数间的协整分析（如美国股票市场是否是巴西股票市场的 Granger 原因，反之亦然），或者是基于股票指数与股票未来价格指数之间的协整关系，或者是基于其他宏观经济变量如通货膨胀率、利率、消费 - 财富比率等之间的关系。这样的列举可以无穷无尽。

一般而言，这些都是纯粹的统计关系，理论证据非常薄弱（或非常简单）。应提醒读者的是，这种可预测性不适于事前分析，交易成本和交易策略风险的考虑都是至关重要的。不幸的是，尽管文献浩如烟海，但我们却无法得出一般结论。但是，很有可能出现的情形是，数据所发现的任何“免费午餐”结果都比投资者按照方程所给出的事前预测进行交易都要昂贵一些。后面将考察小心谨慎的“择时交易”策略，包括样本外预测的绩效。

协整与持续性

P_t 和 D_t 之间通常不会出现协整关系，这是否意味着股票价格的理性价值评估公式（RVF）是错误的呢？我们在这里用蒙特卡罗模拟来对这个问题进行考察。如果贴现因子为常数，那么股票价格由 $P_t = kD_t$ 来决定，这就是 RVF（见第 10 章）。如果 P_t 和 D_t 是 I（1）序列，那么 RVF 意味着 P_t 和 D_t 之间应具有协整关系。经验研究经常发现（P_t，D_t）和（$\ln P_t$，$\ln D_t$）之间都没有协整关系，只有当预期收益为常数时，我们才能由此拒绝 *RVF*。根据构造，*RVF* 是成立的，通过在红利和预期收益具有时变性的模型中使用 MCS，我们能看到为什么 P 和 D 在有限样本中不会出现协整关系。如果在我们所拟合的数据中，标准检验意味着 P 和 D 之间具有协整关系，那么即使 RVF 成立，时变收益也为（从数据中所发现的）无协整关系提供了一种可能的解释。红利和收益合理的数据生成过程为：

$$\ln D_t = \mu + \ln D_{t-1} + \varepsilon_{dt} \quad \varepsilon_{dt} \sim niid\ (0,\ \sigma_d^2) \tag{21}$$

$$(R_t - \bar{R}) = \rho\ (R_{t-1} - \bar{R}) + \varepsilon_{rt} \quad \varepsilon_{rt} \sim niid\ (0,\ \sigma_R^2) \tag{22}$$

因此，$E_t D_{t+m} = \exp\ [m\ (\mu + \sigma_d^2)\ /2]$，在 $\bar{R}$ 附近进行线性化处理后（Poterba 和 Summers（1986）），“理性预期价格”由如下函数来决定：

$$\hat{P}_t = f\ (g,\ \bar{R},\ \rho,\ R_t - \bar{R})\ D_t \tag{23}$$

其中，$g = \exp\ [\mu + \sigma_d^2/2] - 1$。Timmermann（1996）对（21）和（22）使用 MCS，他用（23）来生成 $\hat{P}_t$，然后对（$\hat{P}_t$，D_t）或（$\ln\hat{P}_t$，$\ln D_t$）进行协整检验。结果发现，当预期收益有很强的持续性（即 ρ 大于 0.6）且 RVF 成立（根据构造）时，协整关系会频繁地遭到拒绝（对红利波动率的绝大部分合理值而言）。因此，如果预期收益具有时变性，那么拒绝协整关系并不必然意味着拒绝 RVF。

4.6　非线性模型

在上述 ECM 模型中，非均衡的（$d - p - k$）对 Δp_t 的效应是对称的，且与非均衡的程度独立。非线性模型对这些假设进行了放松。非线性模型在一般情况下显得有些特别，经济理论只是在长期静态均衡的决定中起着重要作用，动态关系由对这种长期均衡的某些非线性反应来决定。非线性模型也包括混沌模型，混沌模型有自己的类别，因为它们都是确定性模型，非线性产生了随机反应。ARCH 和 GARCH 模型的二阶矩（方差和协方差）都具有非线性特征，参

见第 29 章。我们在这里对非线性动态关系建模的一组特别方法进行集中考察。

非线性模型的目的在于放松动态行为假设，这些动态行为就是上述 ECM 模型中所发现的针对非均衡的对称比例反应（保留长期均衡关系）。首先讨论实证研究中大量使用的三种非线性模型：阈值模型（threshold model）、样条模型（spline model）和（二次）平滑转换自回归（smooth transition autoregressive，STAR）模型。这些模型在应用研究中使用得越来越普遍。值得注意的是，它们只能处理动态关系，也通常将长期均衡视为给定。模型通常用非常简单的长期均衡关系（如 $p-d$ 在长期中为常数）来对数据进行经典刻画。

阈值（非对称）模型

阈值模型是建立在“转换”虚拟变量基础上的，模型取决于非均衡项 z_{t-1} 是否高于或低于某个临界值：

$$\begin{aligned}\Delta p_t = &\beta_1(L)\ \Delta d_{t-1} + \beta_2(L)\ \Delta p_{t-1} + \\ &\alpha_1\ (DV_1 = 1,\ z_{t-1} > c_1)\ z_{t-1} - \\ &\alpha_2\ (DV_2 = 1,\ z_{t-1} < c_2)\ z_{t-1} + \varepsilon_t\end{aligned} \tag{24}$$

其中，DV_i 是指标变量，当 z_{t-1} 的条件满足时，它取值为 1（否则就取 0）。如果 $c_1 \neq c_2$，那么对位于 c_1 与 c_2 之间的那些 z_{t-1} 而言，不会有任何调整。以 c_1 和 c_2 上的格子搜索为条件，可以用 OLS 来对模型进行估计。显然，如果 $\alpha_1 = \alpha_2$，那么模型具有对称效应。

样条模型

样条模型是（24）的一个特例，取 $c_1 = c_2 = 0$ 就是这种情形，因此模型不存在阈值，但调整可能是非对称的：

$$\begin{aligned}\Delta p_t = &\beta_1(L)\ \Delta d_{t-1} + \beta_2(L)\ \Delta p_{t-1} - \\ &\alpha_1\ (DV_1 = 1,\ z_{t-1} > 0)\ z_{t-1} \\ &\alpha_2\ (DV_2 = 1,\ z_{t-1} < 0)\ z_{t-1} + \varepsilon_t\end{aligned} \tag{25}$$

STAR 模型

在这种模型变体中，调整的速度不同（$\alpha_1 \neq \alpha_2$），但对非均衡的正负反应不对称：

$$\begin{aligned}\Delta p_t = &\beta_1(L)\ \Delta d_{t-1} + \beta_2(L)\ \Delta p_{t-1} - \alpha_1 z_{t-1}\ [1 - F\ (z_{t-1})] - \\ &\alpha_2 z_{t-1}\ [F\ (z_{t-1})] + \varepsilon_t\end{aligned} \tag{26a}$$

$$F\ (z_{t-1}) = [1 + \exp\{-\gamma\ (z_{t-1} - c_1)\ (z_{t-1} - c_2)\ /\sigma_z^2\}^{-1}] \qquad (\gamma > 0) \tag{26b}$$

如果 $\gamma = 0$，模型是线性的，当 $\gamma \to \infty$ 时，模型向阈值模型靠近。模型对 γ 的中间值而言，对非均衡的调整非常平缓。

ESTAR 模型

比上述模型更简单的一个版本（假设为 *AR*（1）过程）是指数 STAR（ESTAR）模型：

$$z_t = (\pi_0 + \pi_1 z_{t-1}) + (\pi_1^* z_{t-1}) F(z_{t-j}) + \varepsilon_t \tag{27}$$

其中，$F(z_{t-j}) = 1 - \exp[-\gamma(z_{t-j} - c)^2/\sigma_z^2]$。

σ_z^2 项通常取样本均值。$F\ (z_{t-1})$ 为“U 形”，介于 0 与 1 之间。如果 $z_{t-j} = c$，那么 $F = 0$，（24）就变成线性的 AR（1）模型。当 $(z_{t-j} - c) \to \infty$ 时，非均衡的程度很高，且 $F = 1$，因此（27）变成一个不同的 *AR*（1）模型，z_{t-1} 的系数变为 $(\pi_1 + \pi_1^*)$。

从本质来看，ESTAR 模型允许 $z_{t-1} - c$ 的调整速度大大地不同于 0。就股票而言，如果人们相信 $z = d - p$ 主要通过价格变化进行调整，那么这就变成了所谓的风险套利假说（Chiang、Davidson和 Okunev（1997））。一种替代情形是，当 $z_{t-1} - c$ 很大时，调整更加缓慢，因为在等待更早时期错误定价出现反转的过程中，交易者可能会损失如此多的钱，以至于他们只有更少的（借的）资金来对大的非均衡进行套利（Shleifer 和 Vishny（1997）），这就是套利极限假

说。这两个假说可以表示为：

$$\pi_1 + \pi_1^* < \pi_1 \text{ “风险套利”；}$$
$$\pi_1 + \pi_1^* > \pi_1, \text{“套利极限”} \tag{28}$$

对这类模型进行估计时，实践中会遇到几个问题：

（ⅰ）如何确定数据是否存在某种形式的非线性；

（ⅱ）如何选择动态关系中的滞后长度（即 Δ 项），如何选择 F（y_{t-j}）的滞后长度；

（ⅲ）估计和收敛问题；

（ⅳ）模型评估和检验问题。

有好几个预检验程序可以用来确定是否存在一般的非线性，如 BDS 检验、Ramsey 重置检验及其他检验方法（参见 Granger 和 Terasvirirta（1993），Lin 和 Terasvirta（1994））。对 z_t 的非线性进行逼近的流行方法是使用 Taylor 级数。例如，给定 ESTAR 模型（24），可以使用如下方程来进行预检验：

$$z_t = \pi_0 + \pi_1 z_{t-1} + \pi_2 z_{t-2} z_{t-j} + \pi_3 z_{t-1} z_{t-j}^2 + \pi_4 z_{t-1} z_{t-j}^3 + \varepsilon_t \tag{29}$$

于是，对线性与非线性的检验就是对 $\pi_i = 0$（$i = 2-4$）的检验。在 $\hat{z}_t$ 中，Ramsey 检验是类似的，线性模型预测是对更高的幂次方 $\hat{z}_t^q$（$q = 2$，3，…）进行回归，可以用标准 Wald 检验或 LR 检验来考察这些项的统计显著性。上面的叙述也不是没有任何遗漏。例如人们可以用神经网络对 z_t 进行模拟，然后对线性性进行检验（Lee，White 和 Granger（1993）），Chen（2003）提供了对替代性 STAR 模型（即介于 LSTAR 与 ESTAR 之间的模型）进行区分的检验方法（当 $T > 100$ 时，该检验具有很好的功效性质）。滞后长度通常用试错法来确定（如使用 z_t 的偏自相关函数）。

通常用 NLLS 来对上述模型进行估计，尽管有时会遇到收敛问题。STAR 模型中的协整检验（如 p 和 d 之间）和假设检验不是很直接，但在非线性关系的零假设下，要用蒙特卡罗模拟来确定临界值（如 Corradi、Swanson 和 White（2000））。

这些非线性动态模型可以很容易地用于许多问题的研究，在这些问题中，长期均衡方程（如购买力平价、期限结构的预期假说）相当直截了当，我们在这里只对价格红利比的情形进行说明。正如我们所知道的那样，（对数）价格红利比的持续性非常强，在样本均值附近的偏差也较大。ESTAR 可用于对 $z_t = (d-p)_t$ 的调整是否更快、这些偏差是否更大等问题进行考察。Gallagher 和 Taylor（2001）使用美国 1926—1997 年的季度数据发现 AR（1）-ESTAR 模型（F 函数中，$c = 0$，$j = 5$）的 $R^2 = 0.70$，$\gamma = -0.36$（$t = 3.3$）在统计上显著，风险套利命题成立，见（27）和（28）。正如作者所注意到的那样，线性模型的 R^2 为 0.68，因此改进空间并不是很大。但是，模型并不意味着（$d-p$）与均值的距离越大，红利价格比向均值方向的运动速度就越快。然而，模型并不知道进行绝大部分调整的到底是 Δp 还是 Δd，因为ESTAR 模型中的因变量是 $\Delta(d-p)_t$。

显然，这些非线性模型包含了持续性非常强的一些变量（即平稳—非平稳边界线上的变量），危险在于这些结果很有可能是由数据中的一些异常值或区制（状态）转移（regime shift）所造成的。例如，我们知道，在 1929 年大萧条、20 世纪 40 年代、1974 年 5 月或最近的 1987 年等股灾之后，股票价格上升的速度非常快，因此价格红利比也一样。了解这些模型“在整个样本”上的结果对这些“特殊时期”到底有多敏感是一件非常有趣的事情。众所周知，1926—1950 年期间价格红利比的均值比 1950—1997 年期间的均值要高，ESTAR 模型（假设均值为常数）中所发现的非线性效应很有可能是线性动态模型中状态转移所造成的。我们需要对这些模型进行更为深入的研究，尽管“内心的感受”是，在一些时期上或在一些情形下，非线性调整可能会出现。

4.7 马尔可夫转换模型

自 Hamilton（1990）的学术研究开始，马尔可夫转换模型在宏观经济学和金融学的文献中大量出现。马尔可夫转换模型是一种非线性模型，它允许我们将数据视为是由两个或多个体制或“状态”（如“高”和“低”红利价格比或利率状态）所生成的，但这些状态不可观测。因此，在非线性 ESTAR 类型模型中，我们有一个常系数模型，模型对非均衡项的动态反应都是非线性的（但“真实”状态只有一个）。通过马尔可夫转换方法，模型在每种状态下对参数而言都是线性的，但是当你在不同状态之间进行转换时，状态转换的随机本质说明模型中的反应是非线性随机动态反应。从本质来看，两种方法中的条件都比标准 ECM/VECM 常参数模型少，他们只为模型参数简单地提供更多的替代方法，因此也就使模型的适用范围更为宽泛（松散地说，拟合数据有更多的“自由”）。

为了说明问题，我们只考虑两种状态的情形（出于计算困难方面的考虑，应用研究中所考虑的状态数目至多为三个）。马尔可夫性质意味着落在区间 $\{a, b\}$ 内的观察变量 y_t 的概率分布只依赖于 $t-1$ 时刻的状态变量，不依赖于更早时刻的状态。假设不可观察的状态变量为 S_t，因此两状态马尔可夫过程可以表示为：

$$pr\ (S_t = 1 \mid S_{t-1} = 1) = p_{11}$$

$$pr\ (S_t = 2 \mid S_{t-1} = 1) = 1 - p_{11}$$

$$pr\ (S_t = 2 \mid S_{t-1} = 2) = p_{22}$$

$$pr\ (S_t = 1 \mid S_{t-1} = 2) = 1 - p_{22}$$

p_{11}是当前状态为 1 下一个时刻状态仍然为 1 的概率。可以证明，y_t 序列可以表示为：

$$y_t = \mu_1 + \mu_2 S_t + (\sigma_t^2 + \theta S_t)^{1/2} \varepsilon_t$$

其中，$\varepsilon_t \sim N\ (0,\ 1)$。因此，在状态 1 下，均值和方差为 $(\mu_1,\ \sigma_1^2)$；状态 2 下，均值和方差为 $(\mu_1 + \mu_2,\ \sigma_1^2 + \theta)$。用极大似然法估计的未知参数是 $(\mu_1,\ \mu_2,\ \sigma_1^2,\ \sigma_2^2,\ p_{11},\ p_{22})$。转移概率矩阵 P 是由上述概率组成的 2×2 矩阵。当前状态的概率定义为 $\pi_1 = [\pi_{1t},\ \pi_{2t}]$，$\pi_{it}$是 y 在当前时刻所处状态为 i 的概率。给定 π_t 和 2×2 概率矩阵 P，y 在下一期为某个给定状态的概率可由 $\pi_{t+1} = \pi_t P$ 来计算，如此类推有 $\pi_{t+k} = \pi_t P^k$。

Brooks 和 Persand（2001）提供了区制转换模型的一个有趣例子，y_t 是英国 1975 年 1 月—1998 年 8 月期间的 GEYR。他们注意到 GEYR 序列存在值“高”或“低”的时期（与上面所注意到的情形一样），哈密尔顿滤波给出的参数值

$$(\mu_1,\ \mu_2,\ \sigma_1^2,\ \sigma_2^2,\ p_{11},\ p_{22}) = (2.43,\ 2.07,\ 0.062,\ 0.014,\ 0.9547,\ 0.9719)$$

在统计上显著。因此，两种状态中的均值和方差都不同。此外，从“低”GEYR 均值状态向“高”GEYR 均值状态转移的可能性低于 10%。在时间序列中，这说明 GEYR 本身在一个持续（较长）时期中要么呆在低值状态，要么呆在高值状态。

GEYR = $(P_t/B_t)\ (C/D_t)$ 的运动由股票价格 P_t 和债券价格 B_t 所控制。从本质来看，低 GEYR 意味着股票价格相对于债券价格（和红利）而言较低。基于均值反转的交易策略在股票价格低的时候应是购买股票（出售债券）。从本质来看，如果 GEYR 低（高），那么就购买股票（债券）。（在后面的价值增长型选股策略中将看到这一点。）Brooks 和 Persand（2001）使用的就是这种交易策略，但从本质来看，他们没有将 GEYR 视为择时交易的“信号”，而使用的是 GEYR 在下一个时期落到一个特殊状态的预测概率（见后）。因此，如果下一个时期 GEYR 处于低状态的概率大于 0.5，那么就意味着应该购买股票（否则就持有黄金）。状态 1 的预测概率（Engel 和 Hamilton（1990））为

$$p^f_{1,t+1|t} = \mu_2 + [\lambda + (p_{11} + p_{22} - 1)(p_{1t} - \lambda)(\mu_1 - \mu_2)]$$

其中，p_{1t}是所观察到的历史上位于状态1的概率。针对212个数据点进行递推回归（recursive regression），对$p^f_{1,t+1|t}$进行一期预测，我们就可以使用股票－债券的转换交易策略。这与针对股票的“买而持有”被动型交易策略形成对比。1980—1996年期间被动交易策略和转换交易策略的月平均收益分别为0.622%和0.689%，标准差分别为2.14和1.58，夏普比分别为0.089（被动交易策略）和0.16（转换交易策略）。但转换交易策略涉及16次“转换”，考虑交易成本后，被动交易策略和转换交易策略的夏普比几乎相等。对转换交易策略而言，美国和德国的结果要差一些。与简单的“固定参数”时间序列模型（如随机游走、线性AR、MA(1)－GARCH等）相比，（正如所预期的那样）马尔可夫转换模型在样本内统计绩效方面更胜一筹，而在样本外预测绩效方面，模型却相对更为糟糕一些，这是因为模型对下一个时期状态的预测很糟糕（Dacco和Satchell（1999））。

带马尔可夫转换的VECM模型（MS－VECM）

原则上，我们可以将马尔可夫转换（MS）过程嵌入到所有常参数模型中去，因此，如果对标准VECM模型进行扩展，那么就可以将MS－VECM模型写成：

$$\Delta y_t = \mu\ (S_t) + \sum_{d=1}^{p-1} \psi_d\ (S_t)\ \Delta y_{t-d} + \alpha\ (S_t)\ z_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim ND\ (0,\ \Omega\ (S_t)) \qquad S_t = \text{状态}\ 1,\ 2,\ \cdots$$

其中，y_t是由变量组成的$n \times 1$向量（其中的一个变量是股票价格p_t），S_t是马尔可夫转换变量，z_t是协整变量。例如，z_t简单地就为对数价格－红利比率$p_t - d_t$，或标准普尔500指数、金融时报100指数和日经225指数的一个稳定线性组合，或几个国家的远期价格减即期股票指数后的对数$f_t - s_t$（即“基准”，如标准普尔500指数、金融时报100指数和日经225指数的即期－远期基准所构成的3个协整向量$f_t - s_t$）。上面提到的所有z_{t-1}变量都有可能是股票指数变化值Δs_t的Granger原因（也是所选择的协整向量中其他变量变化值的Granger原因）。

现在，实证研究中已经得到很好证实的一个事实是，对价格变化的绝大多数金融时间序列而言（如股票收益、债券收益和外汇收益），Δp_t（即ε_t项）的条件分布不是正态分布，通常具有尖峰、肥尾和左偏的特征。对这些非正态性进行建模的一种方法是使用时变条件方差（和协方差），我们将在第29章讨论这些ARCH－GARCH模型。一种替代方法是，与前面的ESTAR类模型或带马尔可夫转换的VECM模型（MS－VECM）一样使用一些非线性函数，这类模型中的截距项μ、自回归动态演变ψ_d、误差修正系数α和误差的协方差矩阵都依赖于状态。这种“弹性”更强的模型能解释为什么Δy_t（特别地$\Delta p_t \subset \Delta y_t$）在样本内和样本外预测的使用方面及整个分布的形状预测（即预测密度函数）方面都比标准VECM的效果好。

MS－VECM模型的一个很好的例子是Sarno和Valente（2004），他们的模型使用的是1989—2000年的每周数据，模型中的协整向量为$z_{it} = f_{it} - s_{it}$，其中$\{s_i,\ f_i\}$ ＝｛股票指数，期货价格指数｝，i＝标准普尔500指数、日经225指数和金融时报100指数。他们发现，从绝对误差（MAE）和均方根（root mean square）误差来看，MS－VECM的样本外一步预测（one step ahead forecast）并没有对标准VECM模型有所改进。然而，在样本外一步预测中，就正确预测股票价格变化的符号而言，MS－VECM模型的效果比线性VECM模型要好一些——这是对择时交易的一次检验。

因为我们知道协方差矩阵和区制转换VECM模型参数的随机轨道，所以我们能一周接一周地产生股票价格变化样本外预测的预报密度（predictive density）。MS－VECM模型的一步“预测密度”可以用解析的方法来进行计算，密度函数是多变量正态分布的一个组合，组合的权重就是预测的状态概率，主要是后者具有偏度、峰度和异方差方面的特征。

注意，这里的“预报”（predictive）并不是预测的意思，而是指所选择的样本外时期的密度只由（即不是整个样本期）MS－VECM模型给出。于是，在每一个预测期中，“模型”的预报密度可以与由数据得来的真实密度进行比较。Sarno和Valente已经证明，与线性VECM模型

相比，就三种股票指数而言，MS - VECM 模型所给出的“样本外”密度更接近于股票价格变化的真实密度。（看到这一结论被证实是一件很高兴的事情，但结论并不令人感到惊讶，因为 MS - VECM 模型的“设计”目的就是为了拟合数据中已知的非正态性质，而线性 VECM 却不是这样。）

因此，如果有人对股票价格变化 Δs_t 最后的尾部（即 1% 或以下的水平）感兴趣，那么 MS - VECM 模型能更好地测度这一尾部概率下的股票价格变化（因此，在这一显著性水平下，模型对 Var 进行了更好的测度——见 28 章）。

注意 Sarno - Valente 论文中没有说到的那些地方是很重要的。论文没有对市场效率进行检验，只是对相对而言的市场择时交易进行了考察（即与替代性模型相比，他们的模型正确得到 Δs_t 的符号的次数要多一些）。论文没有考虑交易成本和基于 MS - VECM 方法的投资策略风险。就样本内而言，MS - VECM 方法 $\bar{R}^2$ 的取值范围是从日经 225 指数收益的 0.08（8%）到金融时报 100 指数收益的 0.12（12%）（周数据）。因此，在样本外，尽管 MS - VECM 方法在对变化（方向）的预测方面比线性 VECM 有所改进，这个模型还是不能为我们提供模型正确预测 Δs_t 符号的次数方面的信息。最后要注意的是，MS - VECM 方法在 Var 上的卓越表现是针对 1% 置信水平的资产而言的。在实践中，金融机构对多种资产上的潜在损失感兴趣，在计算 Var 时，他们依赖于投资组合中资产之间的误差消除。（他们也经常对 5% 尾部概率下的 Var 感兴趣。）因此，在实践术语中，是投资组合 Var 重要，而不是单个资产的 Var 重要，我们并不知道 MS - VECM 模型（应用于多种资产）是否能对一些简单的方法（如简单的 EWMA）进行改进。我们所知道的是，MS - VECM 模型比第 28 章所讨论的其他方法更加耗费计算机（和人力）资源，因此对 *Var* 的讨论更多的是出于智力上的好奇。当然这些模型大体上都是理论模型，可以视为数据的简单描述。模型仅有的一点“金融”知识是，无风险套利保证了 $f_t = s_t$（如果我们忽略这一关系中剩余的确定性时间 $T - t$ 和随机无风险利率情形，见 Cuthbertson 和 Nitzsche（2001b））。

4.8 交易策略可以获利吗?

当回顾回归模型来试图对收益进行解释时，计量经济学家感兴趣的是一般性的诊断检验（如误差项是否是 *iid*，方程右边的变量是否是弱外生的）、方程在样本外的预测绩效和参数的短期稳定性。在前面的许多研究中，我们并不总是完整地报告这些统计信息，因为很难确定这些结果是否是看上去的那样“稳健”。然而 Pesaran 和 Timmermann（1994）对股票收益早期工作中所面临的批评进行了研究。他们将超额收益对 t 时刻及之前就已知的变量进行回归，但对信息集合的时间处理非常小心。例如，在解释从 1 月末到 1 月末的年收益（使用每年 1 月最后一个交易日的数据）过程中，他们使用直到最后一个交易日的利率（或期限结构变量）数据，但所使用的工业产出数据只是到上一年度的 12 月为止（因为数据公开有一个滞后）。

他们对 1954—1971 年期间及子期间上标准普尔 500 指数和道琼斯工业指数的年度、季度和月度超额收益数据进行了考察。就年度超额收益而言，少数几个变量，包括红利收益、年度通货膨胀率、三个月的利率变化和期限结构，解释了超额收益 60% 的变化。就季度和月度数据而言，宽泛的类似变量分别解释了超额收益 20% 和 10% 的变化。非常有趣的是，在季度和月度回归中，他们发现前面的超额收益对当前收益存在非线性效应。例如，前期超额收益的平方在统计上通常显著，而过去的正收益和负收益对未来收益的影响不一样。作者也给出了序列相关性、异方差、正态性和“正确”函数形式的诊断性检验，这些检验统计量表明模型不存在错误设定。

为了检验这些方程的预测功效，他们使用截止到 t 时刻的数据进行递推估计（OLS），在估计系数的基础上对下期（即）超额收益的符号进行预测。就年度收益而言，70% ~ 80% 的

预测收益有正确的符号，而对季度超额收益而言，回归方程给出的收益预测中，正确符号仍占 65%。于是，Pesaran 和 Timmermann（1994）进一步为早期的超额收益可预测结论提供了证据，他们用少数几个变量就能对结果进行很好的解释。

竞买竞卖利差会产生交易成本（即交易者以较低的价格购买股票，以较高的价格售卖给其他投资者），“买”或“卖”订单都要向经纪人支付佣金。Pesaran 和 Timmermann 使用的“收盘价格”要么是“竞买”价格要么是“竞卖”价格。于是，他们假设，固定的每一美元的买/卖交易成本足以代表所有的交易成本，股票的交易成本 c_s 比债券的交易成本 c_b 要高。他们考虑了一个简单的交易规则，即：

如果所预测的超额收益为正（根据递推回归），那么就持有市场投资组合；否则就持有政府债券，政府债券的期限与交易时期（如年、季度、月）相同。

注意，对这种策略而言，是收益预测的“符号”重要，而不是方程（样本内）的整体“拟合”效果（即 R^2）重要。

上述“转换策略”避免了可能的破产，因为不可能卖空资产，也不存在负债（借款）。被动的基准策略是在任何时候都持有市场投资组合。他们用交易成本是“低”、“中”或“高”来对转换策略和被动策略进行评价。（他们将股票的值设为 0、0.5% 和 1%，将债券的 c_b 值设为 0 和 0.1%。）

就一般情形而言，他们发现，即使交易成本很“高”，转换策略的年收益也比被动策略的年收益要高（即在每年的 1 月份转换一次）（表 5）。然而只有在股票交易成本低于 0.5% 的条件下，转换策略才会对季度或月度区间上的交易支付费用。另外，他们发现转换投资组合年收益的标准差（表 5）比被动投资组合低（即使在交易成本很高的情形下）。于是，在 1960—1990 的整个数据时期上，在均值方差准则下，转换策略占优于被动策略。

表 5　　相对于市场投资组合和国库券而言的标准普尔 500
转换投资组合的绩效测度（年收益，1960—1990）

	投资组合							
	市场[1]			转换[2]			国库券[3]	
交易成本（%）								
股票	0.0	0.5	1.0	0.0	0.5	1.0	–	–
国库券	–	–	–	0.0	0.1	0.1	0.0	0.1
收益和绩效								
算术平均收益（%）	10.780	10.720	10.670	12.700	13.430	12.210	6.750	6.640
收益的标准差（%）	13.090	13.090	13.090	7.240	7.200	7.160	2.820	2.820
夏普比	0.310	0.300	0.300	0.820	0.790	0.760	–	–
Treynor 比率	0.040	0.040	0.039	0.089	0.085	0.081	–	–
Jensen α	–	–	–	0.045	0.043	0.041	–	–
	–	–	–	(4.63)	(4.42)	(4.25)	–	–
期末财富[4]	1 913	1 884	1 885	3 883	3 559	3 346	749	726

1. “市场”投资组合表示买而持有标准普尔 500 指数的投资策略。

2. 转换投资组合的基础是三月利率、期限溢价、通货膨胀率和红利收益等变化的超额收益递归回归。转换规则假设投资组合选择发生在每年 1 月份的最后一个交易日。

3. 国库券表示 12 个月国库券的滚动投资策略。

4. 1960 年 1 月的初始资金为 100 美元。

资料来源：Pesaran 和 Timmermann（1994），复印得到 J. Wiley and Sons 公司的授权。

相对于超额收益方程的不同回归集合和 1960—1970 年、1970—1980 年、1980—1990 年的子样本期而言，上述结果是稳健的。在表 5 中，我们报告了一年期转换投资组合和被动投资组合的夏普比、Treynor 和 Jensen 绩效指数（见第 7 章）。对任意投资组合“p”而言，这些指标为：

$$SR = (ER_p - r)/\sigma_p$$

$$TR = (ER_p - r)/\beta_p$$

$$(R_p - r)_t = \alpha + \beta(R_m - r)_t + \varepsilon_t$$

“Jensen 指数”是上述回归方程的截距 α。这三个统计量是“风险校正后收益”的替代性指标。SR、TR 或 α 越大，投资策略就越成功。我们可以计算转换投资组合、市场投资组合的 SR 和 TR 的平均值。一般而言，除了高成本情景下的月度交易策略之外，Pesaran 和 Timmermann 发现，绩效指标意味着转换投资组合“击败了”被动市场投资组合。

模型的不确定性

Pesaran 和 Timmermann（1995，2000）将上述分析扩展到包括模型不确定性（model uncertainty）的情形。与 1992 年的研究不同，他们允许投资者在任何时点选择“最好”的模型。为了获得最好的模型，三个“核心”变量（红利价格比、利率和通货膨胀率）始终要包含进来，还要对其他变量（利率变化、债券收益率的变化和 1 月虚拟变量）进行检验。此外，如果残差中异常值大于三倍标准差，那么需要在预测方程中加入虚拟变量（此后一直保留这个变量）——这适用于 1974 和 1987 年的股灾时期。

他们使用多个替代性准则来为每个时点选择“最好的”模型（如 $\bar{R}^2$、Akaike 和 Schwartz 准则、“样本内”正确的“符号”预测）。t 时刻转换策略与前面的研究一样，即如果所预测的超额收益 $R_{t+1} - r_t > 0$ 就购买股票指数，否则就持有无风险资产。“被动策略”就是在任何时刻都持有市场指数。

对美国 1960（11）~1992（2）和英国 1965（1）~1993（21）的月度数据进行递推回归，转换策略（对两者都有效，但 Schwartz 选择准则只对美国数据有效）的夏普比比持有被动的市场投资组合的夏普比高。例如，就英国（1970—1993）而言，被动的市场投资组合具有 2.8%的年平均收益，标准差为 36.5%，夏普比为 0.33。针对 Akaike、Schwartz 或 R^2 准测而言，“高”交易成本下的“月度”转换投资组合所具有的平均年收益（标准差）分别为 15.8（10.2）、18.8（11.5）和 15.5（10.5），隐含的夏普比分别为 0.69、0.86 和 0.64（Pesaran 和 Timmermann（2000），表 5）。这三个转换准则的 Jensen α（t 统计量）分别为 0.061（2.8）、0.090（3.62）和 0.058（2.58）。在 20 世纪 80 年代，当市场的波动率低于 70 年代时，转换投资组合给出了更低的绩效结果，但仍然击败了被动的市场投资组合。当主动策略使用的是指数期货而不是购买实际股票时，结果也非常类似。值得注意的是，转换投资策略在英国 1974 年的熊市中并没有“输”，但在 1975 年 1 ~2 月的急剧上升中也没有“赢”。这可能是因为转换策略的在整体上是成功的。

敏锐的赌博者可能会注意到，上述结果意味着“你不得不为你的免费午餐挣钱”。在某些年份中，转换策略也会输钱（如就英国数据而言，不加入 1945—1975 的虚拟变量时，整个样本期上的 R^2 仅为 0.12），你必须拥有足够的资本（或借贷便利）来“承受打击”，在其他日子里艰难地进行交易。当然，每个月你都要进行大量的回归，并期望所发现的任何可预测性都是真实的。就美国而言，Pesaran 和 Timmermann（2000）发现，在 1960—1992 的整个期中没有任何简单模型会表现得足够好，在考虑交易成本之后，在主动（月度）转换策略上很难找到真正的卓越表现（如夏普比）。

人们已经证明（Campell 和 Viceira（1999），Kandel 和 Stamburgh（1996）），多期框架中的投资组合最优化问题严重地依赖于收益的可预测性，当投资者使用择时交易策略时，最优投资组合的权重随着时间会发生急剧变化。Pesaran 和 Timmermann 的证据表明，这样的“主动型”

投资组合策略是值得实施的。

4.9　小结

* 收益可预测性的单变量检验（如自回归系数、自回归模型、ARMA模型、以回归方程为基础的检验、方差比检验等）表明，对总量股票指数而言，短期内（至多6个月）存在正的自相关性但证据较弱，长期内（3～8年的时期）存在均值反转现象且证据较强。不同检验方法所得到的结果也不相同，部分原因是检验统计量的小样本性质不同，也可能是因为这些单变量的性质随着时间的演变不是固定的。单变量检验的功效不是特别的好。

* “择时交易”的多变量检验意味着（实际）股票收益和超额收益是可预测的。这些可预测性再一次受到许多不确定性的约束，在样本外预测检验中有时会“消失”。

* 具有非线性动态演变的模型或嵌入状态变化的模型或递归更新参数的模型都意味着非常弱的样本外可预测性存在，但马尔可夫转换模型考虑了股票收益条件分布的非正态性质。

* 有证据表明，基于回归方程预测结果的主动型交易策略（择时交易）可能会导致经过（事后）风险和交易成本校正后的收益，这个收益只是在某些特定的数据时期中高于被动型交易策略的收益。

* 有效性或EMH检验的核心问题是，在用交易成本和事前风险进行校正后，一连串“赌博”的利润是否为正。在事后结果的基础上，有确凿的证据表明事实上发生的就是这种情形，尽管人们总是认为，用于投资组合风险校正的方法不是很充分。

* 本章关注的主要是收益的统计性质，我们并没有真正提及金融理论。我们已经发现，个体和总量股票价格（收益）具有较高的波动率。在第10章和11章中，我们将考察建立在“基础变量”基础上的各种“理性”经济模型是否能解释这一高波动率现象。如果理论不能解释这一现象，那么我们说我们有过高的波动率（excess volatility）或波动率之谜（volatility puzzle）。

* 类似地，观察到的高股权溢价（及高波动率）和低利率变化应该能由“经济的基础变量”（这是影响投资者的风险来源）来进行解释。许多“标准模型”都不能解释这两个经典事实，这就是众所周知的股权溢价之谜和无风险利率之谜，我们将在第13和14章中将对这些问题进行深入讨论。

* 我们也已经注意到短期收益的正自相关性和长期收益的负自相关性（如均值反转），我们的理论也要对这些现象进行解释。为了发展能解释所有这些典型事实的一般模型，待做事情的菜单非常的长。在一般的模型中，理性的信息灵通的交易者在一定的约束下追求“福利”最大化（如你的消费不可能大于你的终身财富）。

* 第18章中会出现一些惊人的“事实”，即众所知周的异像（anomality），我们期望我们的理论能对此进行解释。就这一点而言，金融经济学家有点像物理学家，他们试图寻找一个很一般的模型来将各种经验结果都包含进去。在以后的章节中，我们将看到他们到底取得了多大的成功。

* 第5～6章的静态模型和第15～16章的跨期模型将讨论资产组合配置的收益可预测性和波动率的含义。

第5章

均值方差组合理论与资本资产定价模型

目的

* 说明在期望收益和风险之间进行权衡的投资者如何持有风险资产的多样化组合。这就得到了有效前沿。
* 确定个体投资者以无风险利率借入（或贷出）多少资金来增加（或减少）“一组”风险资产的风险暴露，这部分由资本市场线（CML）来决定。资本市场线和有效前沿允许我们计算持有风险资产的最优比例，这就是投资组合选择的均值方差模型，这些最优的投资比例就构成了市场投资组合。
* 分析单个证券的均衡期望收益的决定因素，目的是为了使“均值方差”投资者愿意持有所有的风险资产。这就产生了资本资产定价模型（CAPM），模型中的风险溢价取决于资产的“β”。

5.1 概述

我们假设代表性投资者可以选择一组风险资产（股票）和一种无风险资产（例如固定期限的银行存款或三个月期的国库券），只要他们愿意，代表性投资者可以以无风险利率借入或贷出任意金额的资金。假设代表性投资者喜欢更高的期望收益而讨厌风险（即投资者属于风险厌恶型）。将单个证券的期望收益记为 ER_i，假设单个证券的风险由收益的方差 σ_i^2 来度量。所有个体对各种证券的期望收益、方差、协方差（相关系数）都有相同的预期，假设交易成本和税收为0。

考虑持有包含一组风险资产的多样化投资组合的原因。假设投资于无风险资产的资金数量已经固定。将所有财富投资于资产 i 可以获得期望收益 ER_i，同时承担 σ_i^2 的风险。假设现实世界只存在两种资产，且两种资产收益的协方差为负，$\sigma_{12} < 0$，即当第一种资产的收益增加时，第二种资产的收益将减少（这也意味着相关系数 $\rho_{12} = \sigma_{12}/\sigma_1\sigma_2$ 为负）。如果进行分散化投资，同时持有这两种资产，那么你就可以降低整个投资组合（即资产 1 + 资产 2）的方差。为了进一步简化问题，假设 $ER_1 = ER_2$，$\sigma_1^2 = \sigma_2^2$ 且第一种资产的收益每增加1%第二种资产的收益就减少1%（因此，$\rho_{12} = -1$）。在这些条件下，将初始财富平均投资于两种资产，整个投资组合的期望收益就等于 ER_1（$= ER_2$）。然而，分散投资将投资组合的风险降低到0：当第一种资产的期望收益超过平均水平时，第二种资产的期望收益就会低于平均水平（因为 $\rho_{12} = -1$）。这当然是一个特例。一般来说，资产收益之间的协方差即使为0或正（但不是完全正相关），持有两种资产的组合仍然可以分散风险。

上述简单例子说明，每个投资者在允许以无风险利率无限制地借入或者贷出资金的情况下至少持有几只股票的原因。为了说明这一点，我们举一个反例。如果一支股票最初没有使任何投资者满意，那么当投资者售卖股票时，这只股票的当前价格会下降。然而，在其他条件相同的情形下，当前价格的下降表明下一期的期望收益会更高（假设投资者期望在将来支付红利）。因此，人们会看到当前价格的下降一直到预期收益上涨，从而股票变得非常有吸引力以至于人们愿意持有。

现在，也许读者正在考虑，如何在某个时点上分析个体投资者的感觉或偏好。然而，存在一个非常著名的结论，这就是分离原理。投资决策可以考虑两个单独的决策，第一个决策是风险资产最优持有比例 ω_i^* 的选择，这一决策与个体投资者在风险和收益之间进行权衡的主观偏好相独立，选择仅依赖于个体投资者对客观市场变量如预期收益、方差和协方差等的看法。假设所有投资者关于这些变量的预期都是同质的，所有个体都持有相同比例的风险资产（例如，所有投资者都持有 1/20 的“股票 *A*” 和 1/80 的“股票 *B*”），持有比例与投资者的风险偏好无关。因此，所有投资者都将持有与（总的）市场投资组合权重相同的风险资产（例如，如果 AT&T 的股票价值占总股票市场指数的 1/20，那么所有投资者都将其风险资产组合的 1/20 投资于 AT&T 股票）。

在决定了最优市场投资组合以后，个体投资者的偏好才会进入到计算过程中来。在决策过程的第二阶段，个体投资者要决定借入（或贷出）多少资金来增加（或减少）（以固定比例）投资于市场组合风险资产的初始财富数量，就是在这一点上个体投资者的偏好才会进入决策过程。如果个体投资者是风险规避程度很高的投资者，他会将绝大部分财富投资于无风险资产（获得收益率 r），而将少部分财富以固定的比例 ω_i^* 投资于风险资产。相反，对风险规避程度不是很高的投资者而言，他将以无风险利率借入的资金（和初始财富一起）以最优比例投资于固定的一组风险资产。然而，要注意的是，涉及个体偏好的第二阶段不会与风险资产的相对需求（即投资比例 ω_i^*）产生矛盾。因为风险资产的相对需求决定了均衡的预期收益，所以后者与个体投资者的偏好无关（只取决于如收益的方差和协方差等客观市场变量）。

在这一章和随后的章节中，我们将使用如下等价的表示方法表示期望收益、方差和协方差：

$$\text{期望收益} = \mu_i \equiv ER_i$$

$$\text{收益的方差} = \sigma_i^2 \equiv Var\ (R_i)$$

$$\text{收益的协方差} = \sigma_{ij} \equiv Cov\ (R_i,\ R_j)$$

资本资产定价模型（CAPM）与 β

现在转为讨论由 CAPM 产生的与均衡收益有关的一个特殊结果。CAPM 提供了一种完美的方法来决定市场中任何单个风险资产的均衡期望收益 ER_{it}。单个风险资产的期望超额收益 $(ER_i - r)_t$ 和市场组合的期望超额收益 $(ER_m - r)_t$ 直接相关，固定的投资比例由个体风险资产的 β 来确定：

$$(ER_i - r)_t = \beta_i\ (ER_m - r)_t \tag{1}$$

其中

$$\beta_i = \text{Cov}\ (R_{it},\ R_{mt})\ /\text{Var}\ (R_{mt}) \tag{2}$$

当市场的期望超额收益给定后，CAPM 模型可以解释资产 i 的期望超额收益。因此，CAPM 模型不能解释市场收益。为此，我们需要一个更精确的模型（参见第 13 章的随机折现因子模型）。由于自变量和因变量都是 t 时刻的数据，所以 CAPM 模型不是对资产 i 收益的预测方程。但是，CAPM 模型意味着，资产 i 超额收益 $(ER_i - r)_t$ 的运动与同期的市场超额收益变化相关。

ER_{mt}是市场投资组合的期望收益，也是以最优投资比例 ω_i^* 持有所有资产的“平均”收益。由于市场组合的实际收益和期望收益之间存在差异，所以市场组合收益的方差 Var (R_{mt})

不为0。β_i 的定义表明资产 i 的均衡期望收益取决于：

（ⅰ）证券和市场投资组合之间的协方差 Cov（R_{it}，R_{mt}）；

（ⅱ）与市场投资组合收益的方差 Var（R_{mt}）成反比例关系。

概括地说，如果能够计算出事后（或实际）收益的平均值、事前期望收益 ER_{it} 的近似值，那么我们可以将 CAPM 模型解释为证券 i 的（比方说几个月）平均收益。

为了以一种直观的方式来阐述 CAPM，考虑如下问题。资产 A 收益的标准差为20%，β 值为1.5，资产 B 收益的标准差为60%，β 值为0.75。考虑只持有一种资产的情形，A 或者 B（没有其他风险资产），问自己希望哪种资产能获得更高的平均收益。资产 B 比资产 A 有更大的标准差，所以推测你希望资产 B 有较高的平均收益（相对于 A 而言），因为资产 B 有更高的（个体）风险。你的预感将是正确的。

现在考虑一个稍微不同的问题。假设你已经持有很多风险资产（比方说25），你想在现有投资组合中增加 A 或者 B，那么，增加哪种资产你会获得更高的平均收益呢？A 的 β 值为1.5，B 的 β 值为0.75，β 值与（A 或者 B）资产收益和你所持有的现有资产组合收益之间的协方差（相关系数）成比例。如果你将这种资产纳入到你的投资组合中去，那么这种资产的 β 值越低，这种相关性就越低，投资组合风险（标准差）就增加得越少。因此，资产 B 所导致的投资组合风险的增加要小一些（相对于 A 而言），所以即使 B 有相对较低的平均收益，你也希望把它增加到你的投资组合中去。因此，在决定平均收益时，是增加的投资组合风险重要，而不是个体资产的风险（即标准差）重要。这就是投资组合理论与 CAPM 模型在概念上的联系：当证券作为多样化组合的一个部分而被持有时，β 值为什么能成为证券风险一种正确度量。

这里还有另外一种直观的观点。资产 B 收益的标准差为60%、β 值为0.75，当作为多样化投资组合的一部分而被持有时，我们一致认为即使 B 比 A（标准差为20%，β 值为1.5）有较低的期望收益，我们也更愿意持有资产 B。资产 B 的较大标准差是如何影响组合的标准差的呢？当你在一个充分分散化的投资组合中同时持有这种股票和其他股票时，这种股票的绝大部分风险就能以几乎为零的成本被“分散掉”。因此，即使 B 有较高的标准差，你也不会获得更高回报（平均收益）。CAPM 模型表明，在股票 B 对你整个组合的风险作出贡献的基础上你仅能获得一个平均收益，收益的大小由 β 值来决定。

就股票市场上单个证券均衡收益的绝对水平而言，CAPM 模型能告诉我们什么呢？首先，注意到 $ER_m - r > 0$，否则风险厌恶的代表性投资者确定将所有财富投资于无风险资产可以获得更多收益时，就不会持有风险资产的市场组合。

个体股票（同一国家内）的收益倾向于向相同的方向运动，因此，从实证角度来看，Cov（R_{it}，R_{mt}）$\geqslant 0$，$\beta_i > 0$。CAPM 模型预测的结果是，那些与市场投资组合有零协方差的股票，只要期望收益率等于无风险利率就会被持有（（1）中令 $\beta = 0$）。与市场收益有较大的正协方差的证券（$\beta_i > 0$）将会获得相对较高的期望收益，这是因为将这样的证券增加到投资组合中去几乎不会减小整个投资组合的方差。

CAPM 模型也允许人们在个体股票 β_i 值（我们假设这些值的度量是准确的）的基础上来评价股票期望收益的相对波动性。$\beta_i = 1$ 股票的收益和市场投资组合的收益相同（$ER_i = ER_m$），这些股票被称为中性股票。如果 $\beta_i > 1$，那么股票就被称为是“进攻型”股票，因为它的期望收益的变化（上升或下降）比市场投资组合期望收益的变化要大。反之，如果 $\beta_i < 1$，那么就称这些股票为“防守型”股票。因此，投资者可以使用 β 值对各种证券的相对安全性进行排序。然而，后者也不应该从投资组合理论的重要预测中删除掉，即所有投资者都应该以相同的最优权重 ω_i^* 持有股票。投资组合应包括中性股票、进攻型股票和防守型股票。希望持有某种特殊股票的投资者会使用 β 值对组合中所有股票进行排序（即投资者不会遵循 CAPM 模型的假设，因此不会试图模拟市场组合）。个体投资组合（n 种资产）的 β 值为 $\beta_p = \sum_{i=1}^{n} \omega_i \beta_i$。

5.2　均值方差模型

我们假设投资者偏好较高的期望收益 ER，厌恶风险（即属于风险规避类型）。我们用投资组合的方差来衡量风险大小。因此，如果投资者面临组合 A（n 种证券）和组合 B（一组不同的证券），那么根据均值方差准则（MVC），如果

$$ER_A \geqslant ER_B \tag{i}$$

且

$$\mathrm{Var}\ (R_A)\ \leqslant \mathrm{Var}\ (R_B) \quad 或 \quad \mathrm{SD}\ (R_A)\ \leqslant \mathrm{SD}\ (R_B) \tag{ii}$$

那么投资组合 A 要优于投资组合 B，SD 为标准差。例如，如果 $\mathrm{E}R_A > \mathrm{E}R_B$ 但 Var（R_A）> Var(R_B)，那么使用 MVC，我们就不能说对投资者而言哪个组合要更好一些。满足均值方差准则（MVC）的所有组合都被称为有效投资组合。组合 A 比组合 B 有较低的期望收益和较高的方差，我们就认为组合 A 是无效的。如果组合 B 是可获得的，那么（原则上讲）投资者永远不会持有组合 A。

我们希望以一种简单的方式来说明持有分散化资产组合的收益，我们一开始就假设仅有两种（风险）资产，这两种风险资产的（单期）实际收益为 R_1 和 R_2，期望收益分别为 $\mu_1 = \mathrm{E}R_1$，$\mu_2 = \mathrm{E}R_2$。每一种资产收益的方差为 $\sigma_i^2 = E\ (R_i - \mu_i)^2$。假设各种收益之间的相关系数为 ρ（$-1 \leqslant \rho \leqslant 1$），其中

$$\rho = \sigma_{12}/\sigma_1\sigma_2 \quad 且 \quad \sigma_{12} = E\ [\ (R_1 - \mu_1)\ (R_2 - \mu_2)\ ] \tag{3}$$

$\sigma_{12} = \mathrm{Cov}\ (R_1,\ R_2)$ 为两个收益之间的协方差。如果 $\rho = +1$，那么两种资产完全正（线性）相关，资产收益总是朝同一个方向运动。当 $\rho = -1$ 时，就会出现相反的情形；当 $\rho = 0$ 时，资产收益之间就不是（线性）相关的。正如我们在下面要看到的那样，由资产 1 和资产 2 构成的投资组合的“风险大小”在很大程度上取决于相关系数 ρ 的大小和符号。如果 $\rho = -1$，那么将最初财富以一个特定比例投资于两种资产就可以完全消除投资组合的风险。即使 ρ 为正（但 $\rho < 1$），通过分散化，整个组合的风险也会降低（尽管不为 0）。

最小方差“有效”组合

假设投资者选择以一个特定比例将他的总财富投资于各种资产是为了使组合的风险达到最小。在这一阶段，不允许投资者借入或贷出或将他的财富投资于无风险资产。投资者是“把所有鸡蛋放在一个篮子里面”而把所有财富只投资于资产 1 或资产 2，同时承担风险 σ_1 或 σ_2，还是将所有财富分散投资于资产 1 和资产 2？如果是分散化投资，那么每一种资产的投资数量又为多少呢？假设投资者选择以 ω_1 的财富比例投资于资产 1，以 $\omega_2 = 1 - \omega_1$ 的财富比例投资于资产 2。分散化投资组合的实际收益为（一期之后这个数据才会报告出来）：

$$R_p = \omega_1 R_1 + \omega_2 R_2 \tag{4}$$

投资组合的期望收益为（在时期开始的时候就形成这种预期）

$$ER_p = \omega_1 ER_1 + \omega_2 ER_2 = \omega_1\mu_1 + \omega_2\mu_2 \tag{5}$$

投资组合的方差为

$$\begin{aligned} \sigma_p^2 &= E\ (R_p - ER_p)^2 \\ &= E\ [\omega_1\ (R_1 - \mu_1)\ + \omega_2\ (R_2 - \mu_2)\ ]^2 \\ &= \omega_1^2\sigma_1^2 + \omega_2^2\sigma_2^2 + 2\omega_1\omega_2\rho\sigma_1\sigma_2 \end{aligned} \tag{6}$$

我们假设投资者在选择 ω_i 时并不关心期望收益。为了使投资组合的风险 σ_p^2 达到最小，我们将 $\omega_2 = 1 - \omega_1$ 代入，并对未知参数 ω_1 求导：

$$\frac{\partial\sigma_p^2}{\partial\omega_1} = 2\omega_1\sigma_1^2 - 2\ (1 - \omega_1)\ \sigma_2^2 + 2\ (1 - 2\omega_1)\ \rho\sigma_1\sigma_2 = 0 \tag{7}$$

由（7）式解得 ω_1，

$$\omega_1=\frac{(\sigma_2^2-\rho\sigma_1\sigma_2)}{(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)}=\frac{(\sigma_2^2-\sigma_{12})}{(\sigma_1^2+\sigma_2^2-2\sigma_{12})} \tag{8}$$

由（8）式可以注意到，当 $\rho=-1$ 时，投资组合的方差最小，当 $\rho=+1$ 时，投资组合的方差达到最大（假设 $\omega_i>0$）。为了说明问题，假设 $\sigma_1^2=(0.4)^2$，$\sigma_2^2=(0.5)^3$，$\rho=0.25$（即正相关）。于是，使方差最小的 ω_1 的值为：

$$\omega_1=\frac{(0.5)^2-0.25\ (0.4)\ (0.5)}{(0.4)^2+\ (0.5)^2-2\times0.25\times0.4\times0.5}=\frac{20}{31} \tag{9}$$

将 ω_1 代入到（6）式就可以得到，$\sigma_p^2=12.1\%$，这比将所有财富只投资于资产 1 的方差 $\sigma_1^2=(0.4)^2=16\%$ 或投资于资产 2 的方差 $\sigma_2^2=(0.5)^2=25\%$ 都要小。注意，一般而言，最小方差组合有一个正的期望收益，在这里为 $ER_p=(20/31)\ ER_1+(11/31)\ ER_2$，尽管 ω_i 的选择与合意的期望收益相独立。

就 $\rho=-1$ 的特殊情形而言，使用（8）式我们就可以得到 $\omega_1=5/9$，将其代入（6）式就得到 $\sigma_p^2=0$，因此，当两种资产是完全负相关的时候，所有风险都能够被完全分散掉。从上述分析可以看出，单独持有一种资产的风险可能较高（即收益的方差较大），但是如果这种资产与已经持有的投资组合中的资产负相关，那么即使这种资产的期望收益相对较低，投资者也愿意将这种资产增加到现有的组合中去，因为这种资产能够降低整个投资组合的风险 σ_p^2。这一直观观点解释了 CAPM 模型决定资产均衡收益的原因。

保险原理

可以将上述分析推广到（6）式包含 n 种资产的情形，我们有：

$$\sigma_p^2=\sum_{i=1}^{n}\omega_i^2\sigma_i^2+\sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n}\omega_i\omega_j\rho_{ij}\sigma_i\sigma_j$$

就资产收益完全无关（对 n 种资产而言都有 $\rho_{ij}=0$）的情形而言，投资组合的方差可以简化为

$$\sigma_p^2=\ (\omega_1^2\sigma_1^2+\omega_2^2\sigma_2^2+\cdots+\omega_n^2\sigma_n^2) \tag{10}$$

进一步进行简化，如果所有的方差都相等（$\sigma_i^2=\sigma^2$）且每种资产都以相同的比例（$1/n$）持有，那么我们有：

$$\sigma_p^2=\frac{1}{n^2}\cdot n\sigma^2=\frac{1}{n}\sigma^2 \tag{11}$$

因此，当 $n\to\infty$ 时，投资组合的方差接近于 0。于是，如果不相关的风险被合并在一起，那么投资组合风险的绝大部分都能够被分散掉，这就是“保险原理”，也是你们的房屋和汽车保险费相对于它们本身的价值而言很小的一个原因。保险公司在确定保费的时候在很大程度上依赖于如下事实：这些风险在个体之间不相关。与每个证券相联系的总风险中，部分是那些对所有证券都产生影响的因素（如利率、经济增长率等）所引起的——这就是著名的市场风险，部分是由对个体公司而言的特殊事件所引起的（如新订单的运气、投入成本预料不到的下降等）——这就是所谓的特殊（异质型）风险。当我们持有一个由多种股票组成的分散化投资组合时，异质型风险能够被完全消除。从本质上讲，就是“运气较好”的几个公司能够被当前正在经历“坏运气”的其他几个公司完全抵消掉。一般说来，如果你持有多种股票，这些异质型风险是能够被消除掉的。从直觉来看，人们倾向于持有如下观点：持有任何股票的平均收益都不能反映这种异质型风险（假设你将所有的股票放在一个分散化的投资组合中去）。就像我们即将看到的那样，CAPM 模型就承载了这种直觉。

在实践中，即使随机地选择股票（如从标准普尔 500 中进行选择），即使股票之间不存在零相关性，人们也能大大地降低投资组合的方差。当你将所持有股票的数量从 1 增加到 10 时，投资组合的方差 σ_p^2 会下降得非常的快。当投资组合中股票数量超过 30 后，投资组合方差的减少幅度会变得非常的小（图 1）。这个原因及佣金费用和监管大量股票的信息成本就可以解释

为什么个体投资者倾向于投资较小数目的股票组合。个体投资者也可以通过投资共同基金和养老基金来获得分散投资的好处，因为这些机构使用来自大量个体投资者的资金去投资类别非常宽泛的金融资产，每个投资者都拥有这个“大投资组合”的一个部分。注意，我们还没有谈论任何与上述“随机股票选择”策略的期望收益有关的东西，尽管当股票数目接近 30 时，你将获得与 S&*P*500 相同的平均收益。问题是，我们能否在均值方差组合理论的基础上，通过选择（最优的）资产比例来改进这个随机选择。

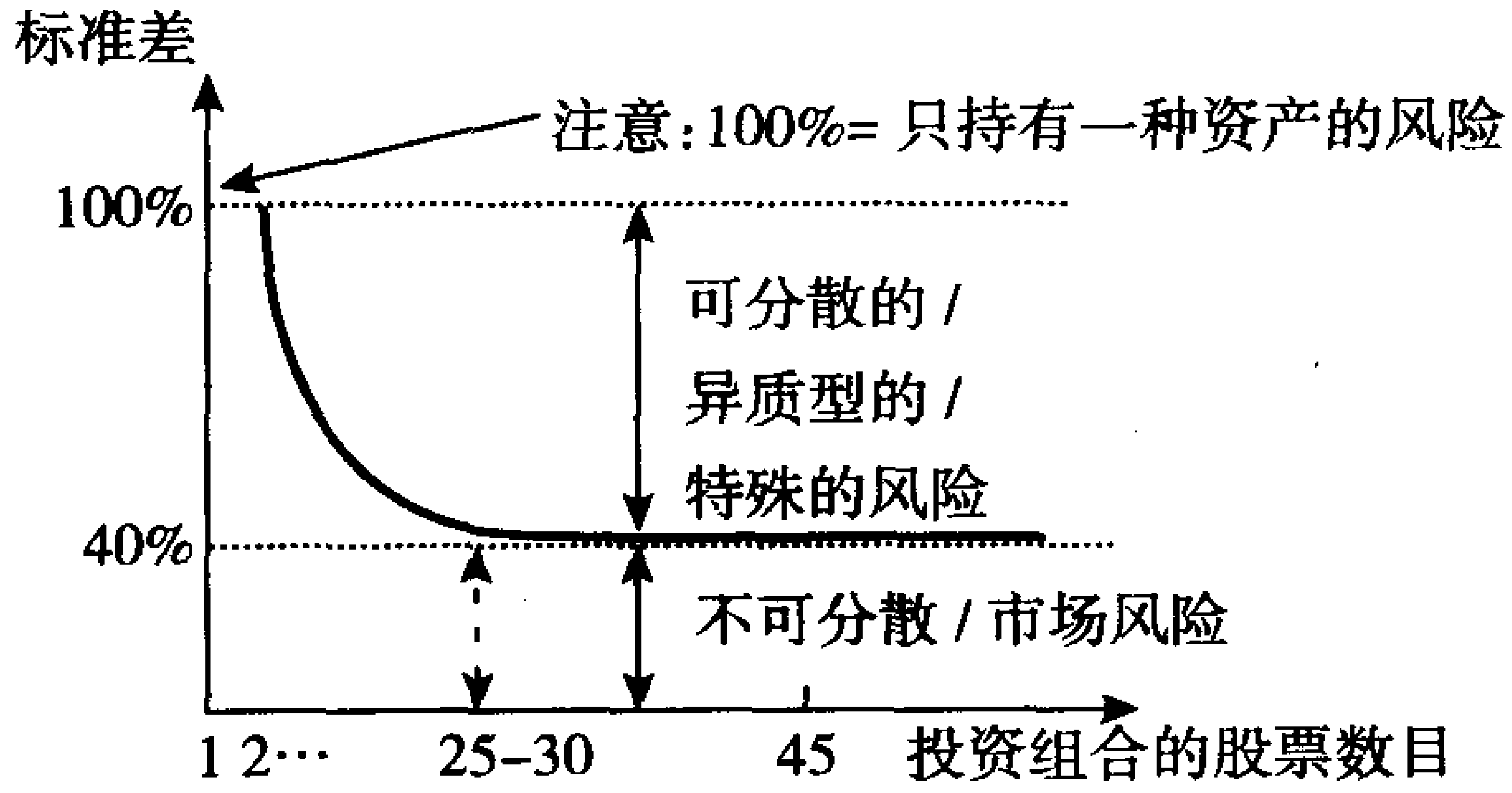


图 1 股票的随机选择

投资组合的期望收益和方差

很明显，个体投资者感兴趣的是投资组合的期望收益 $\mu \equiv ER_p$ 和方差 σ_p。我们现在要问的问题是，当投资者改变投资于每种风险资产上的财富比例时，相对于其他资产而言，μ_p 和 σ_p 如何发生变化。考虑只有两种资产的情形。μ_1，μ_2，σ_1，σ_2（或 ρ）是固定的且是已知的。当我们改变 ω_1（$\omega_2 = 1 - \omega_1$）的大小时，由方程（5）和方程（6）就可以得到与任意选择 ω_1（和 ω_2）相对应的组合（μ_p，σ_p）的值。（注意，这里不存在最大化/最小化问题；这纯粹是一个算术计算问题。）对 $\mu_1 = 10$，$\mu_2 = 20$，$\rho = -0.5$，$\sigma_1 = 10$（$\sigma_1^2 = 100$），$\sigma_2 = 30$（$\sigma_2^2 = 900$）而言，表 1 给出了相应的数值结果，图 2 给出了相应的图形。

表 1 投资组合均值和方差的计算（$\rho = 0.5$）

ω_1	ω_2（$= 1 - \omega_1$）	投资组合的期望收益	投资组合的标准差	投资组合的方差
0	1	20	30	900
0.2	0.8	18	23.07	532
0.4	0.6	16	16.37	268
0.6	0.4	14	10.39	108
0.8	0.2	12	7.21	52
1	0	10	10	100

我们可以对不同的 ρ（$-1 \leq \rho \leq 1$）来重复上面的计算。一般而言，从图 3 可以看出，当 ρ 接近于 -1 时，点（μ_p，σ_p）就向纵轴靠近，这说明 ω_i 对任意给定的期望收益而言，存在进一步降低投资组合风险的可能性。（分别针对 $\rho = 0.5$ 和 $\rho = -0.5$ 来对投资组合 *A* 和 *B* 进行比较。）当 $\rho = -1$ 时，曲线和纵轴相交，这说明取某些值时可以使组合的风险降低到 0；当 $\rho = 1$ 时，“风险—收益”曲线就变成一条直线，并且每个证券的（μ_i，σ_i）都在这条直线上。在上面的例子中，我们任意地选择了一组 ω_i 值，这里并不涉及最大化问题。在现实世界中，（在任何时间点上）ρ 只能有一个取值，因此，与不同 ω_i 值相对应的只有一个“风险—收益”点。

风险收益组合是投资者可行集合或机会集合的一个部分，就两种资产的情形而言，所给出的图形就是著名的风险－收益前沿。

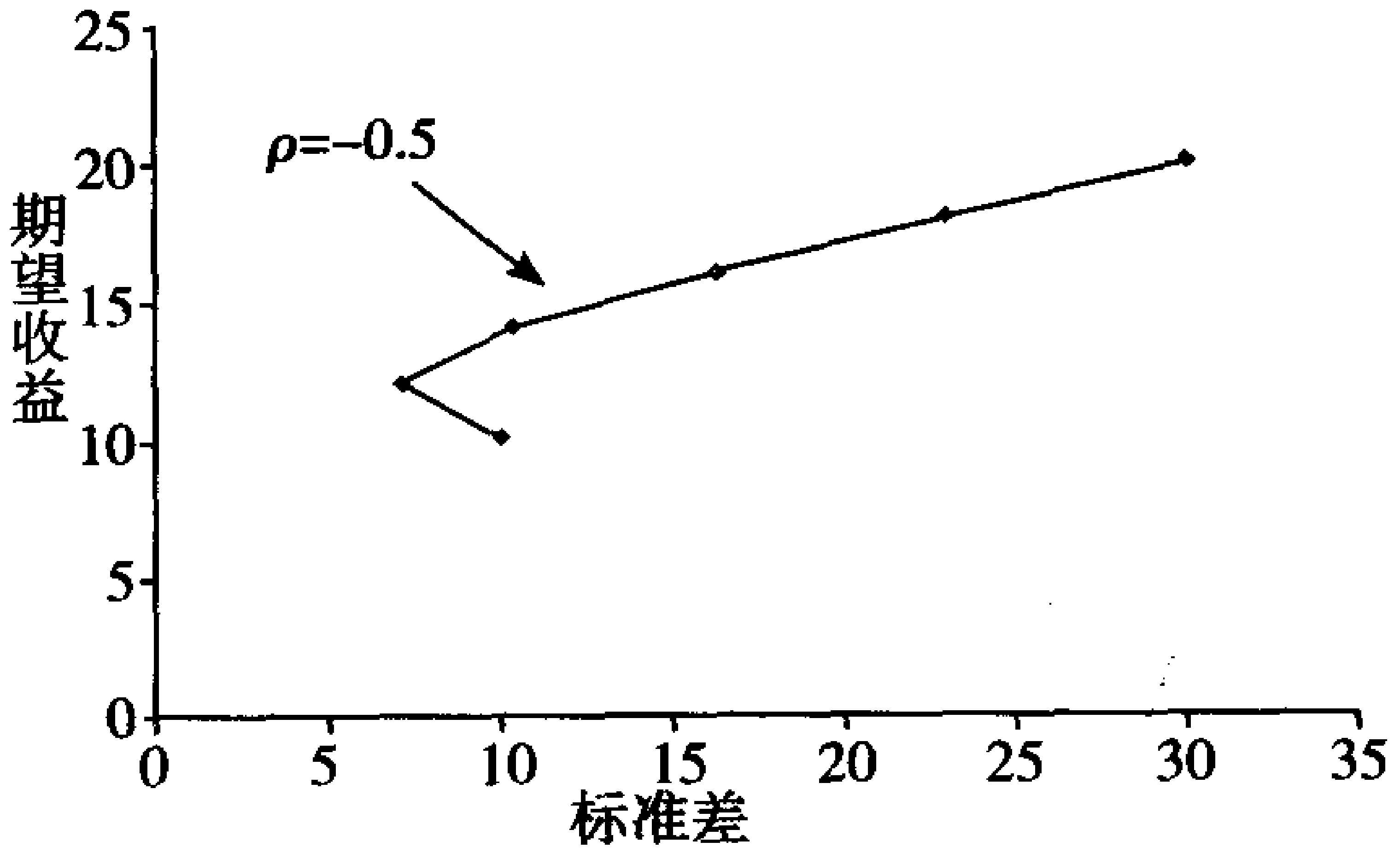


图2 期望收益和标准差

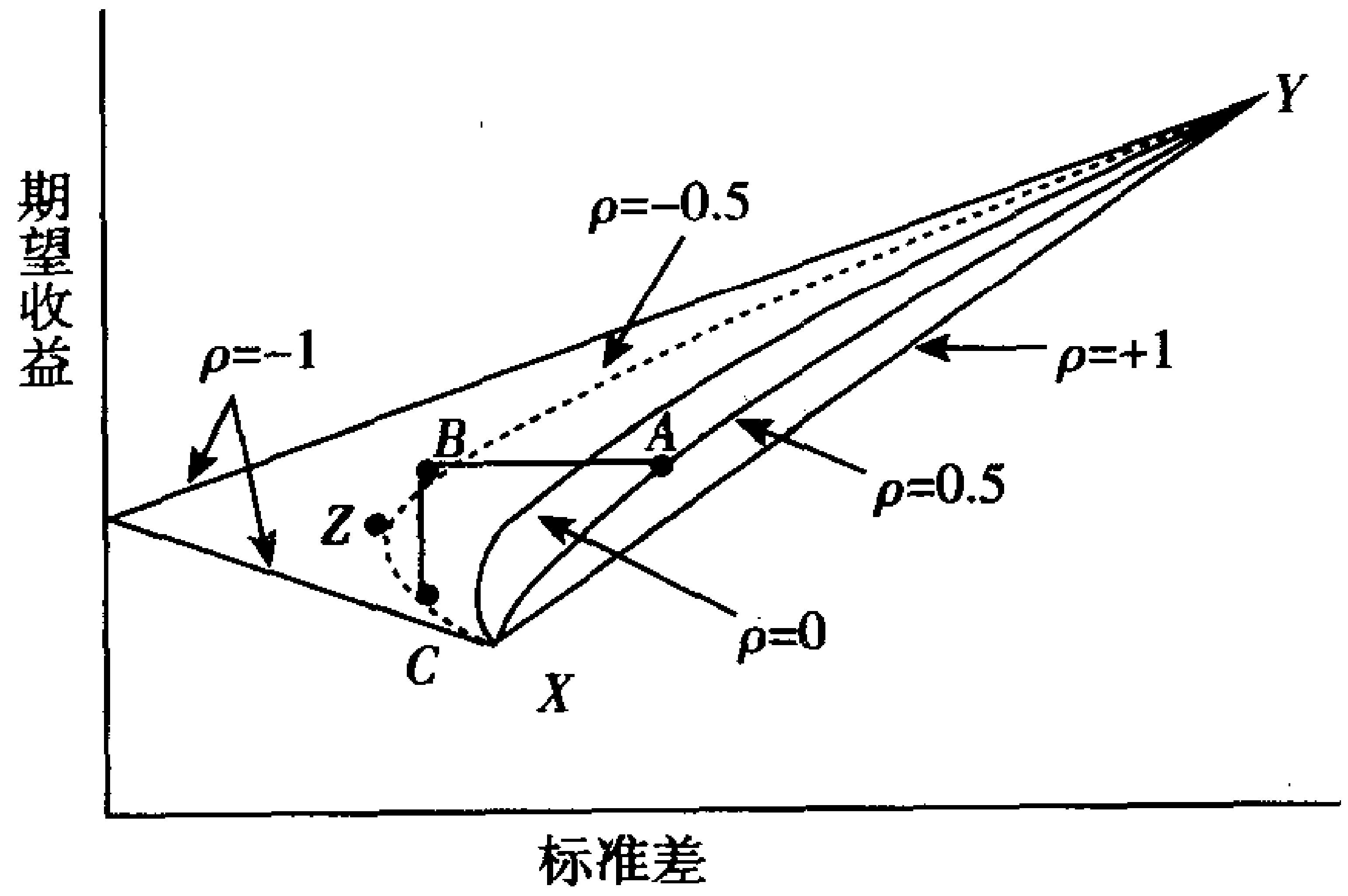


图3 有效前沿和相关性

两种以上的证券

如果我们允许投资者将他的财富在任意两种证券之间进行配置，A 和 B，或 B 和 C，或者 A 和 C，那么我们就得到了三个图形，每一个图形都是“两种资产”的组合。如果现在允许投资者在三种证券上同时进行投资，也就是说，我们可以改变 ω_1、ω_2 和 ω_3 的比例（$\sum_{i=1}^{3}\omega_i=1$），那么这个新的“三种资产的前沿”将位于任何“两种资产前沿”的右边。这说明，在任意给定的期望收益水平下，持有更多证券可以降低投资组合的风险。从直觉来看，根据组合（μ_p，σ_p），毫无疑问，有更多的选择能给你更好的结果。通过改变三种资产的持有比例 ω_i，有效前沿的斜率就可以度量代表性投资者在收益和风险之间的权衡程度。注意，最小方差点 Z 以下的虚线部分（就图3 中的前沿而言，$\rho=-0.5$）是均值—方差无效组合。投资者永远不会选择投资组合 C，而是选择投资组合 B，因为与投资组合 B 相比，投资组合 C 有着相同的风险水平但期望收益更低。在均值方差准则（MVC）下，我们称投资组合 B 占优于投资组合 C。

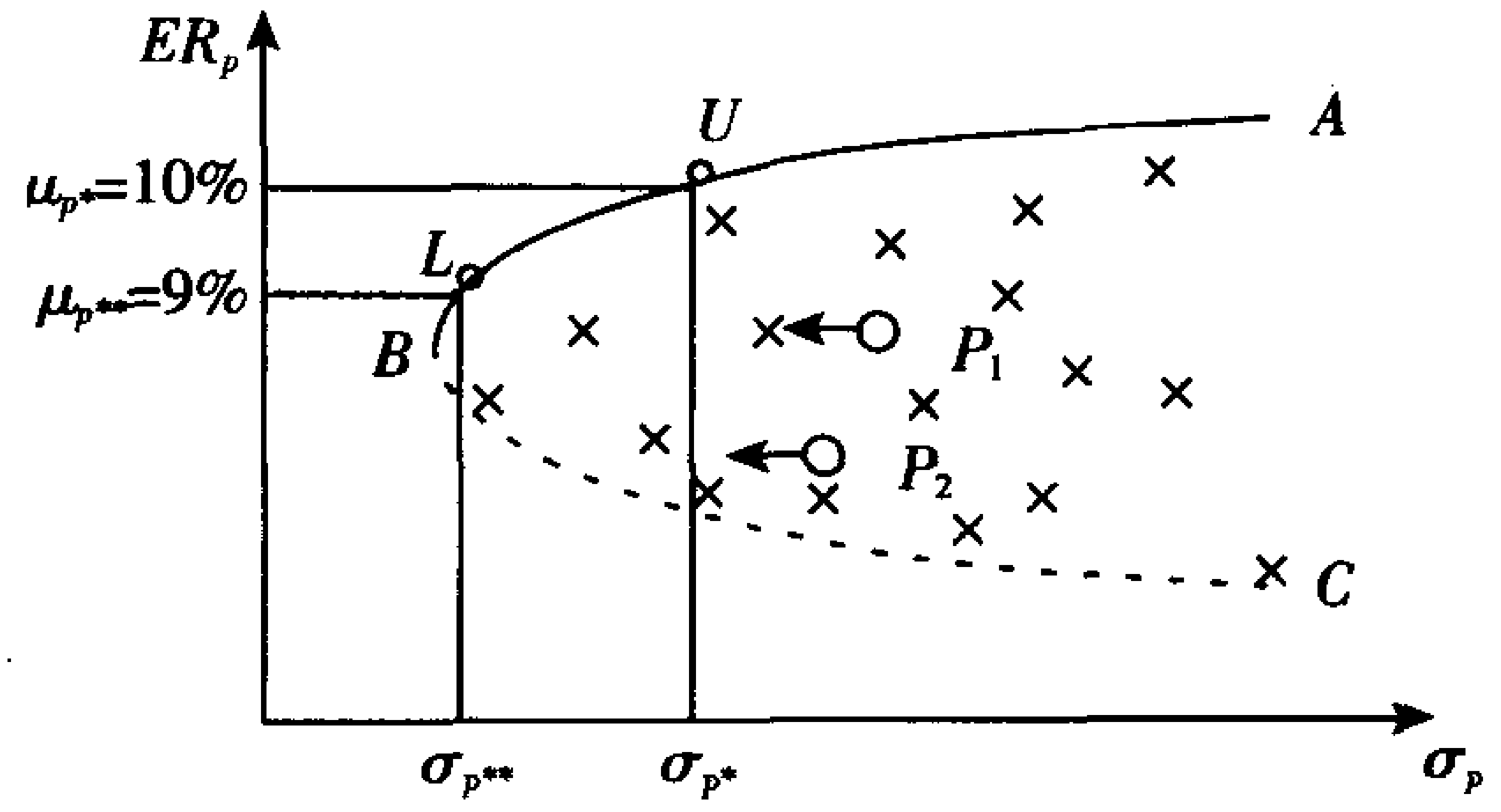


图 4　有效投资组合和无效投资组合

有效前沿

现在，我们考虑 N 种资产的情形。当改变 N 种资产的持有比例 ω_i（$i=1, 2, \cdots, N$）来构造投资组合时，我们可以得到大量可能的投资组合，这与我们构造第 2 种、第 3 种直到第 N 种资产组合的情形一样。我们也可以用相同数目的资产但投资比例不同来构造投资组合。每个可能的投资组合在图 4 中都用“X”来标识。

如果我们现在使用均值方差占优准则，那么曲线 ABC 上的点都占优于投资组合机会集合内部的所有点（如图 4 中的 P_1 和 P_2），因为曲线 ABC 上的点在期望收益相同的条件下有更小的方差。曲线 AB 上的点也占优于曲线 BC 上的点，所以曲线 AB 代表了有效投资组合集合中不同的持有比例 ω_i，这部分曲线被称为有效前沿。

为了构造有效前沿，投资者应怎样计算 ω_i 的值呢？投资者所面临的已知变量有：N 种资产的期望收益 μ_i 和方差 σ_i^2，$n(n-1)/2$ 个协方差 σ_{ij}（或相关系数 ρ_{ij}）。投资组合的期望收益和方差公式为：

$$\mu_p = \sum_{i=1}^{n} \omega_i \mu_i \tag{12}$$

$$\sigma_p^2 = \sum_{i=1}^{n} \omega_i^2 \sigma_i^2 + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} \omega_i \omega_j \rho_{ij} \sigma_i \sigma_j \tag{13}$$

其中，$\sigma_{ij}=\rho_{ij}\sigma_j\sigma_j$。我们假设投资者希望选择每种资产的投资比例 ω_i，但关心的是期望收益和风险。有效前沿表明，(μ_p, σ_p) 的所有组合都是在给定的期望收益 μ_p 下使投资组合风险 σ_p 达到最小的那些组合。投资者的预算约束为 $\sum\omega_i=1$，也就是说，所有财富都用于风险资产的投资。（就目前而言，我们不允许他以无风险利率借入或贷出资金。）允许投资者进行卖空操作，$\omega_i<0$。投资者如何寻找有效前沿的一种典型方法包括如下步骤。

（1）为投资组合选择“目标”收益 μ_p^*（如 $\mu_p^*=10\%$）；

（2）任意选择 μ_p^* 可以达到的财富投资比例 $(\omega_i)_1$（$i=1, 2, \cdots, n$）（使用方程（12））；

（3）使用（13）式，用这些 $(\omega_i)_1$ 值来计算投资组合的方差或标准差 $(\sigma_p)_1$；

（4）如果 $\sigma_{p2}<\sigma_{p1}$，那么就用一组新的 $(\omega_i)_2$ 值来重复第 2 步和第 3 步的计算，然后扔掉 $(\omega_i)_1$，保留 $(\omega_i)_2$；反之亦反之。

（5）重复第 2 步到第 4 步，直到所得到的一组资产比例 ω_i^*（$\sum\omega_i^*=1$）满足目标收益 μ_p^* 且使投资组合的标准差达到最小，记为 σ_p^*。以 ω_i^* 持有的资产组合就是有效投资组合，这就得到了 (μ_p, σ_p) 空间上的一个点，如图 4 中的 U 点；

（6）另外选择一个目标收益 μ_p^{**}（比方说等于 9%），重复以上步骤，就可以得到一个新的有效组合，以 ω_i^{**} 为投资比例，最小标准差为 σ_p^{**}——L 点。

选择不同的目标收益 μ_p 来重复以上步骤，就可以得到曲线 $AULBC$（图 4）。然而，只有曲线的上半部分，即 $AULB$，才是有效投资组合集合，这一部分也被称为有效前沿。

这一阶段值得注意的是，上述问题的解 ω_i^* 在一般情形下（通常是这样）可能为正也可能为负。$\omega_i^*>0$ 表示购买股票（即股票的头寸为“做多”），$\omega_i^*<0$ 表示“卖空”股票，也就是说，投资者从其他拥有股票的人（如经纪人）那里借入股票，然后在市场上将其卖出。因此，她在这些股票上的持有份额为负（他必须在未来的某个时刻将股票还给经纪人），他用卖空股票的资金来增加其他股票的持有数量。总之，有效前沿“涉及”以下几个步骤。

（1）投资者选择最优投资比例 ω_i^*——投资比例满足约束条件 $\sum\omega_i^*=1$，在投资组合的期望收益水平 μ_p 给定的条件下来最小化标准差 σ_p。[1]

（2）投资者重复以上程序，计算每一个期望收益 μ_p 下 σ_p 的最小值，所描绘出的（μ_p，σ_p）图形就是有效前沿。对于一组给定的（μ_i，σ_i，ρ_{ij}）值，我们只有一个有效前沿。

（3）有效前沿上的每一点都对应于一个不同的最优投资组合，最优投资组合中股票的投资比例分别为 ω_1^*，ω_2^*，ω_3^*，…。

以上三点就是投资者使用分离原理时的第一部分“决策”——现在，我们转入决策过程的第二部分。

借入和贷出：变换线

现在，我们允许投资者以无风险利率 r 借入或贷出资金。因为 r 在整个持有期内是固定的，它的方差和（与 n 种风险资产的）协方差都为0。因此，投资者可以：

（ⅰ）将所有财富投资于风险资产，不进行任何借贷；

（ⅱ）将部分财富投资于风险资产，将余下部分以无风险利率贷出；

（ⅲ）以无风险利率借入资金，和总财富一起投资于风险资产。在这种情形中，投资者被认为是持有一个杠杆组合。

变换线是由无风险资产和风险资产所构成投资组合的期望收益和风险之间的关系。对由这两种资产构成的任何组合而言，变换线都成立，可以证明，期望收益和风险（用“新”投资组合的标准差来度量）之间的相互关系是线性关系。假设我们构造了一个投资组合（称它为 k），投资组合由两种资产构成，一种是风险资产，期望收益为 μ_1，标准差为 σ_1，另外一种资产是无风险资产。可以证明，新投资组合收益和标准差之间的关系为：

$$\mu_k=a+b\sigma_k$$

其中“a”和“b”都是常数，μ_k 为新组合的期望收益，σ_k 为新组合的标准差。类似地，我们也可以构造另外一个新组合 N，它由如下资产构成：（ⅰ）以 ω_i 的比例持有 q 种风险资产所构成的风险组合；（ⅱ）无风险资产。我们再一次有如下关系：

$$\mu_N=\delta_0+\delta_1\sigma_N$$

为了导出变换线方程，假设个体投资者已经选择了 q 种风险资产（股票）的一个特殊比例组合，风险资产的实际收益为 R，期望收益为 μ_R，方差为 σ_R^2。注意，ω_i 并不是最优比例，但可以任意取值（约束条件为 $\sum_i\omega_i=1$）。现在，投资者考虑的是，将自己的财富以多大的比例投资于这 q 种风险资产所组成的投资组合，以无风险利率借入或贷出多少资金。于是，考虑一个“新”的投资组合，即由无风险资产和一组风险资产所构成的投资组合。如果将财富的 x 部分投资于无风险资产，将（$1-x$）部分投资于“风险资产”组合。用 R_N 和 μ_N 分别表示新组合的实际收益和期望收益。

$$R_N=xr+(1-x)R \tag{15}$$

$$\mu_N=xr+(1-x)\mu_R \tag{16}$$

其中，（R，μ_R）是投资者所持有投资组合中风险资产部分的（实际收益，期望收益）。当 $x=1$ 时，所有财富都投资于无风险资产，$\mu_N=r$；当 $x=0$ 时，所有财富都投资于风险资产（股票），$\mu_N=\mu_R$；当 $x<0$ 时，代表性投资者以无风险利率 r 借入资金来投资风险资产组合。例如，当 $x=-0.5$ 且初始财富为100美元时，个体投资者（以无风险利率 r）借入50美元，然后将150美元投资于股票（即杠杆头寸）。

因为 r 在整个持有期内是确定的且是已知的，所以“新”组合的标准差只取决于股票风险组合的标准差 σ_R，由（15）和（16），我们有

$$\sigma_N^2 = E\ (R_N - \mu_N)^2 = \ (1-x)^2 E\ (R-\mu_R)^2 \tag{17}$$

$$\sigma_N = \ (1-x)\ \sigma_R \tag{18}$$

σ_R 是风险资产组合的标准差。方程（16）和（18）都是定义方程，但是，根据“新”组合的均值和标准差（μ_N，σ_N），可以将它们变换为更简单的方程，这对后面的分析非常有用。重新排列（18），

$$(1-x)\ = \sigma_N / \sigma_R \tag{19}$$

将（19）中 x 和（$1-x$）代入（16）就得到了一个恒等式

$$\mu_N = r + \left[\frac{\mu_R - r}{\sigma_R}\right] \cdot \sigma_N = \delta_0 + \delta_1 \sigma_N \tag{20}$$

其中，$\delta_0 = r$，$\delta_1 =$ （$\mu_R - r$）$/\sigma_R$。于是，对由两种资产（一种是风险资产（组合））构成的投资组合而言，另外一种资产是无风险资产，新组合的期望收益 μ_N 和 σ_N 标准差之间是线性关系，截距为 r，斜率为 δ_1。当然，方程（20）是一个恒等式，它不涉及任何行为。注意，（$\mu_R - r$）总是为正，否则没有任何人会持有风险资产。

当投资组合仅包含 n 种风险资产时，正如我们已经看到的那样，收益—标准差空间中的有效机会集合是一条曲线（图4）。然而，对由无风险资产和任意单一风险资产所构成的两个投资组合而言，有效机会集合是一条为正的直线。这不会引起混淆，因为两种情形所考虑的投资组合不同，在“有效集合”情形中，曲线是在最优条件下导出来的，而不仅仅是（两个）恒等式的重新排列。

方程（20）说明，μ_N 是关于 σ_N/σ_R 的递增函数。这是因为，根据（19），σ_N/σ_R 的增加意味着，财富中风险资产的投资比例（即 $1-x$）会增加。因为 $\mu_R > r$，所以这就使新投资组合的期望收益 μ_N 增加。类似地，对给定的（σ_N/σ_R）=（$1-x$），风险资产期望超额收益（$\mu_R - r$）的增加使整个投资组合的收益 μ_N 也增加。这仅仅是因为投资者在这里持有的是固定比例（$1-x$）的风险资产，但后者的超额收益更高。

从（20）可以看出，当将所有财富都投资于一组风险资产时，$x=0$，从而 $\sigma_N = \sigma_R$，即持有的是100%的股票组合（图5中的 X 点）。当所有的财富都投资于无风险资产时，$x=1$，$\mu_N = r$（因为 $\sigma_N/\sigma_R = 0$）。在 r 和 X 之间的那些点上，个体投资者将一部分财富投资于无风险资产一部分投资于风险资产。在像 Z 的那些点上，个体投资者持有的是杠杆组合（即他以无风险利率 r 借入资金，并将所有财富投资于股票）。

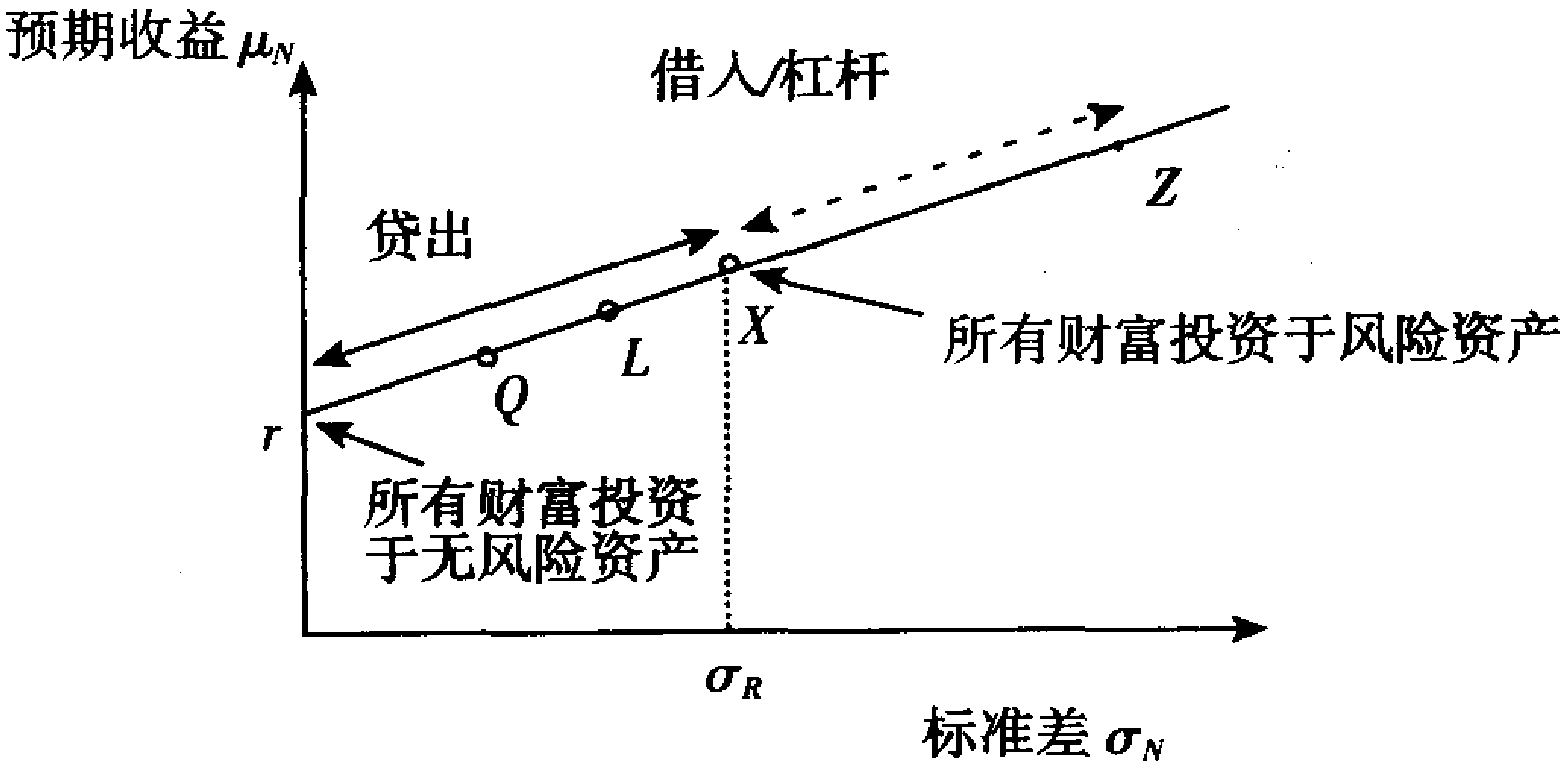


图5 变换线

最优“切点”组合

变换线给出了任何由无风险资产和一组风险资产所构成的投资组合的风险和收益之间的

关系。变换线的导出并不涉及代表性投资者的优化问题：它是一个恒等式。在给定变换线上的任意点上，个体投资者都以相同的比例 ω_i 持有风险资产。假设点 X（图 5）代表的是分别以 20%、25% 和 55% 的比例持有 3 种风险证券“α”、“β” 和 “γ”。Q、L 和 Z 点表示以相同的比例持有风险资产。沿着变换线，唯一的“数量”变化是所持有的风险资产相对于无风险资产的比例。

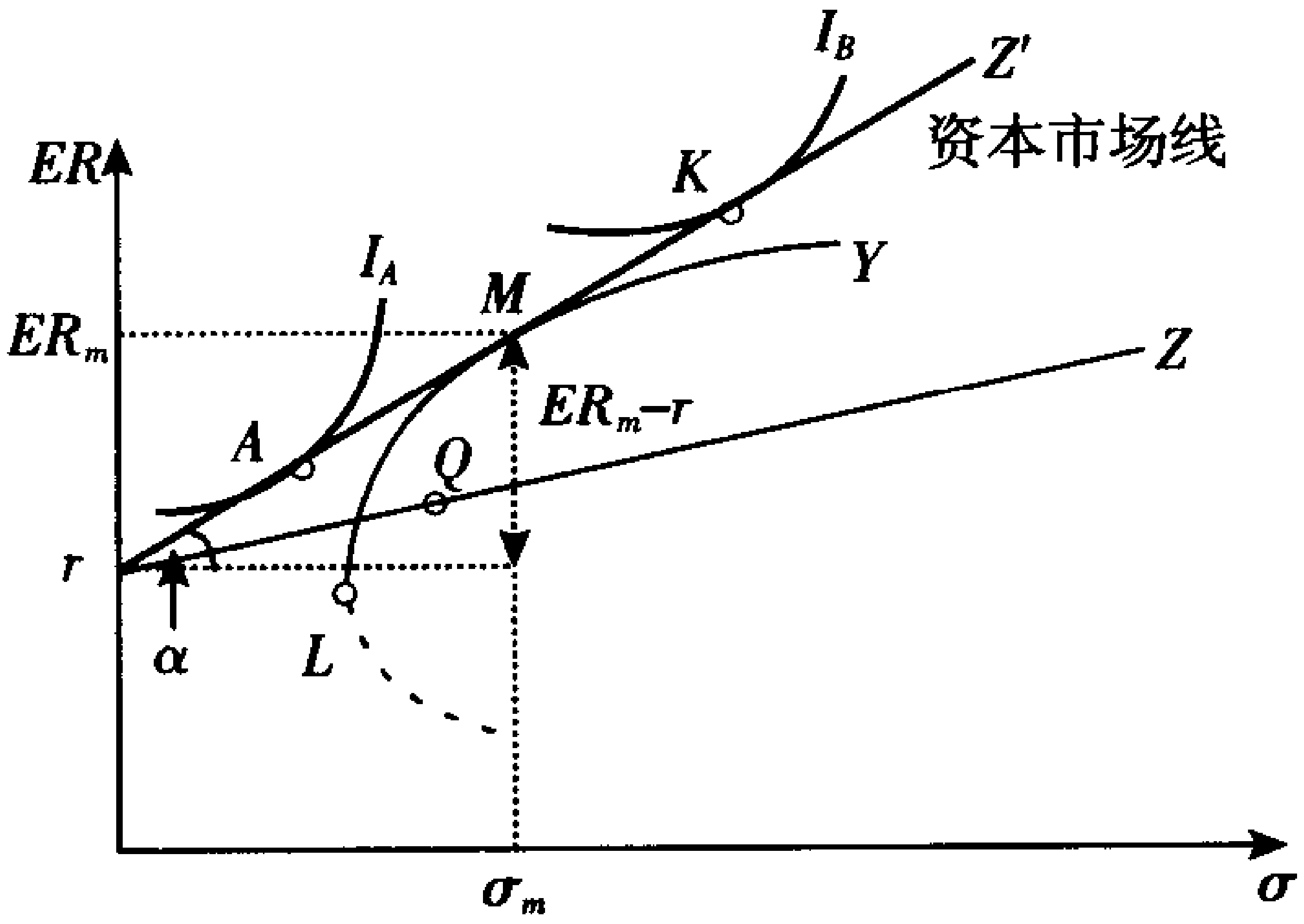


图 6 投资组合选择

沿着变换线 rZ，在任何一点上投资者都可以借入或贷出资金。（准确地说，沿着变换线 rZ，投资者究竟位于哪一点取决于他对风险和收益的偏好，但是，正如我们即将看到的那样，除了在很靠后的章节中，这一考虑并没有进入我们的分析。）例如，图 5 中的 Q 点代表的是，40% 的财富投资于无风险资产，60% 的财富投资于风险资产组合。于是，拥有 100 美元财富的投资者在 Q 点持有 40 美元的无风险资产，持有 60 美元的风险资产：持有 $0.2\times60=12$ 美元的证券 α，持有 $0.25\times60=15$ 美元的证券 β，持有 $0.55\times60=33$ 美元的证券 γ。

尽管投资者能够到达变换线 rZ 上的任意一点，任何一个投资者（无论他的偏好如何）都会更偏好于变换线 rZ'（见图 6）。这是因为在 rZ' 的任意一点上，与 rZ 上的点相比，投资者有着更高的期望收益。事实上，因为 rZ' 与有效前沿相切，所以它为投资者提供了更好的可行机会集合。因为 M 点在有效前沿上，所以以 ω_i 的比例持有风险资产是最优的（即前面提到的 ω_i^*）。投资者可以是 rZ' 上的任意一点，但 M 点始终是所有投资者都持有的股票组合。于是，称 M 点为市场投资组合，称 rZ' 为资本市场线（*CML*）。因此，资本市场线（*CML*）是与有效前沿相切的一条变换线。

投资者的偏好仅仅决定个体投资者会选择资本市场线（*CML*）上的哪一点。例如，投资者仅有一点风险厌恶或没有任何风险厌恶，这样的投资者会选择 K 点，在 K 点他可以（以无风险利率 r）借入资金来增加财富，然后将所有资金投资于 M 点所代表的一组证券（但它仍然以固定的比例 ω_i 持有所有股票）。

沿着资本市场线，每个投资者都会持有相同比例的风险资产，但投资于风险资产上的财富数量不同。如果投资者借入 50 美元，他的资金总额就变为 150 美元，那么他在 α、β 和 γ 三种风险资产上的投资金额分别为 30 美元、37.5 美元和 82.5 美元（$=150\omega_i^*$）。这一杠杆组合的风险为（使用（18））：$\sigma_N=(1-x)\ \sigma_m=1.5\sigma_m$，期望收益为 $\mu_N=1.5\ (\mu_R-r)\ +r$。于是，杠杆组合位于 M 点的“东北”方，如 K 点（图 6）。

分离原理

投资者将作出两个独立的决策：

（ⅰ）投资者使用期望收益、方差和协方差的知识来计算有效前沿上的股票组合，然后确定从 r 出发的一条与有效前沿相切的直线，切点为 M 点。这些操作都不需要考虑投资者自身的

偏好。所有的投资者，无论偏好如何（但对预期收益等的看法相同），都会以 M 点的最优比例 ω_i^* 来持有风险资产。所有投资者都持有市场投资组合，更为准确地说，所有投资者都以风险资产在市场上的相对价值比例来持有风险资产。换一种叙述方式，如果所有投资者都以相同比例持有风险资产，那么市场仅仅是所有投资者的总和，于是从逻辑来看，所有投资者都必须以市场投资组合的比例来持有资产。因此，如果 AT&T 公司的股票价值占股票市场总价值的 10%，那么每个投资者所持有的 AT&T 公司的股票价值也是所持有的风险组合价值的 10%。

（ⅱ）投资者再决定风险资产和无风险资产的持有比例。这个决策依赖于他的主观风险—收益偏好，在 M 点的左边，个体投资者是合理的风险厌恶者，他将财富的一部分投资于市场组合（固定的最优比例 ω_i^*），将另一部分财富投资于无风险资产。如果个体投资者是风险偏好类型，那么他会选择 M 右边的点，如 K 点，他持有的是一个杠杆组合（即通过借钱来增加市场投资组合的持有数量）。在 M 点，投资者将所有财富投资于市场组合，既不需要以无风险利率借入资金也不需要贷出资金。

根据（20），以市场投资组合 M 点为切点的资本市场线 rZ' 必然有如下形式：

$$\mu_N = r + \left[\frac{\mu_m - r}{\sigma_m}\right] \cdot \sigma_N \tag{21}$$

风险的市场价格

个体投资者持有无风险资产和一组风险资产的数量取决于他的感觉或偏好。相对而言的风险厌恶者——无差异曲线为 I_1，将一部分财富投资于无风险资产，而将另外一部分财富投资于市场组合（图6，A 点）。风险偏好者会借入资金来投资于比初始财富还要多的风险资产（图6，K 点）。然而，所有投资者都有一个共同点，最优风险资产组合都位于资本市场线上。对每个投资者而言，

$$\text{资本市场线的斜率} = (\mu_m - r)/\sigma_m = \text{无差异曲线的斜率} \tag{22}$$

资本市场线的斜率也经常被称为风险的市场价格。无差异曲线的斜率被称为边际替代率（MRS），因为它是个体在高风险高收益下“权衡”的比率。

所有投资者的投资组合都位于 CML 上，于是，他们都面临着相同的市场风险。根据方程（20）和图6，很明显，对 A 和 K 处的两个投资者而言，风险的市场价格都等于 MRS。因此，在均衡状态下，所有投资者在风险和收益之间的权衡相同。不仅如此，在资本市场线上的任意一点，“每单位风险的超额收益”相对于有效前沿 LMY 上其他投资组合而言是最大的（图6）。于是，无风险资产的引入增加了投资者的可行选择。

根据衡量风险大小的标准差，可以导出有效前沿和市场组合的偏差。当风险用投资组合的方差来进行度量时，

$$\lambda_m = (\mu_m - r)/\sigma_m^2 \tag{23}$$

λ_m 也经常被称为风险的市场价格。因为 σ_m 和 σ_m^2 在概念上是非常相似的，一般不会引起混淆。（在一般均衡框架下（23）的推导可参见 Merton（1973），当风险用两种不同的方式来进行度量时，CAPM 模型两种表示之间差异的相关讨论可参见 Roll（1977））

以怎样的比例持有资产？

在允许借贷时，我们知道，个体投资者将以 M 点所表示的最优比例来持有一组风险资产。所有投资者都选择用 M 点表示的风险资产比例是因为，在有效集合或一组风险资产给定的条件下，以无风险利率 r 借入或贷出资金可以使投资者到达最高的变换线。（这最终使每个投资者处于最高的那条无差异曲线上。）

但是，仍然存在一个问题。我们应怎样计算用 M 点表示的风险资产持有比例 ω_i^* 呢？到目前为止，我们只说明了如何计算有效前沿上每一点的最优比例 ω_i^*，还没有说明特殊点 M 上的比例 ω_i^* 是如何导出的。为了说明这一点，注意，对变换线上的任何投资组合而言（图

6)，有

$$\tan\alpha = (ER_p - r)/\sigma_p \tag{24}$$

p 表示任意的风险组合，正如我们已经看到的那样，ER_p 和 σ_p 取决于 ω_i（及风险资产 μ_i 和 σ_{ij} 的已知值）。于是，为了到达 M 点，在 $\sum\omega_i = 1$ 的约束条件下，最大化关于 ω_i 的方程（24）就得到了最优投资比例 ω_i^*。有些 ω_i^* 可能小于0，这表示卖空资产。（如果不允许卖空，那么对所有的就必须施加另外的约束条件 $\omega_i^* \geqslant 0$，但这违背了用来推导 CAPM 模型的假设——见后。）

5.3 资本资产定价模型

为了使所有投资者的有效前沿都相同，对基础市场变量 μ_i、σ_i^2 和 ρ_{ij} 而言，所有投资者必须有相同的预期。于是，在相同的预期下，所有投资者都以市场组合 M 点的比例来持有风险资产。同质预期的假设对市场均衡的产生非常重要，均衡时，所有风险资产都被投资者以 M 点的最优比例 ω_i^* 所持有，换句话说，市场是出清的。

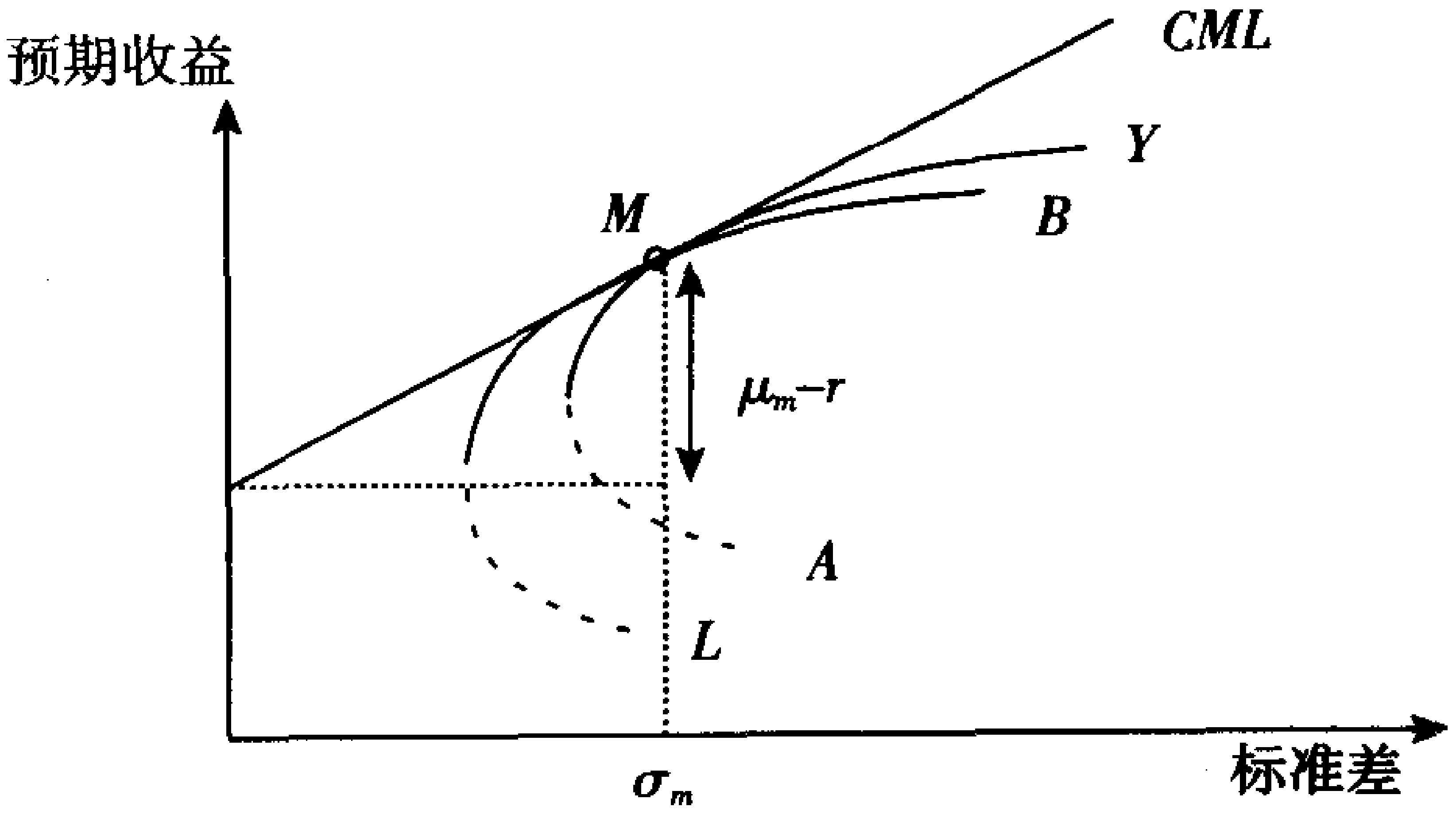


图7　市场投资组合

代表性投资者选择最优资产比例的均值方差模型也就是我们所熟知的 CAPM 均衡预期收益模型（假设预期同质）。均衡收益是个体投资者以相同比例持有最优资产份额的反映，这个比例与个人财富无关。我们在这里说明了，CAPM 模型在只运用图形分析（图7）和简单代数的情况下是如何解决均值方差优化问题的。（我们在下一章会给出一个正规的推导。）

资本市场线（CML）的斜率为常数，它代表的是对所有投资者而言都相同的风险市场价格。

$$CML\text{ 的斜率} = (\mu_m - r)/\sigma_m \tag{25}$$

现在，我们来做个思维实验，从 M 点（它以固定比例持有所有资产）开始“运动”，通过将一些资金投资于 M 点所表示的现有资产，另一些资金投资于任意风险资产 i 来构造一个人工组合。这个人工组合（称为 p）由两个风险组合所构成：以 x_i 的比例投资于资产 i，以 $(1-x_i)$ 的比例投资于 M 点的组合。投资组合 p 的期望收益为 μ_p，标准差为 σ_p：

$$\mu_p = x_i\mu_i + (1-x_i)\mu_m \tag{26a}$$

$$\sigma_p = [x_i^2\sigma_i^2 + (1-x_i)^2\sigma_m^2 + 2x_i(1-x_i)\sigma_{im}]^{1/2} \tag{26b}$$

投资组合 p 位于曲线 AMB 上，在 M 点相切。它不会穿过有效前沿，因为根据定义，有效前沿是在任意给定的期望收益水平下使方差最小的投资组合。也要注意的是，在 M 点，没有借入或贷出。当我们改变比例 x_i，并沿着 MA 移动时，我们就在“卖空”证券 i，并将超过100%的资金投资于组合 M。

推导的关键是要注意，在 *M* 点，曲线 *LMY* 和 *AMB* 恰好相交，因为 *M* 点是市场投资组合，所以 $x_i=0$。为了找到有效前沿在 *M* 点的斜率，我们要求

$$\left[\frac{\partial\mu_p}{\partial\sigma_p}\right]_{x_i=0}=\left[\frac{\partial\mu_p}{\partial x_i}\right]\left[\frac{\partial\sigma_p}{\partial x_i}\right]^{-1}\tag{27}$$

所有导数都是在 $x_i=0$ 处进行计算的。根据（26a）和（26b），

$$\left[\frac{\partial\sigma_p}{\partial x_i}\right]=\mu_i-\mu_m$$

$$\left[\frac{\partial\sigma_p}{\partial x_i}\right]=\frac{1}{2\sigma_p}\ [2x_i\sigma_i^2-2\ (1-x_i)\ \sigma_m^2+2\sigma_{im}-4x_i\sigma_{im}]\tag{28}$$

在 $x_i=0$ 处（*M* 点），我们知道 $\sigma_p=\sigma_m$，于是

$$\left[\frac{\partial\sigma_p}{\partial x_i}\right]_{x_i=0}=\frac{\sigma_{im}-\sigma_m^2}{\sigma_m}\tag{29}$$

将（28）和（29）代入（27），

$$\left[\frac{\partial\mu_p}{\partial\sigma_p}\right]_{x_i=0}=\frac{(\mu_i-\mu_m)\ \sigma_m}{\sigma_{im}-\sigma_m^2}\tag{30}$$

但在 *M* 点，有效前沿的斜率（方程（30））和资本市场线 *CML*（方程（25））的斜率相等，

$$\frac{(\mu_i-\mu_m)\ \sigma_m}{\sigma_{im}-\sigma_m^2}=\frac{\mu_m-r}{\sigma_m}\tag{31}$$

根据（31），我们就得到了 CAPM 的关系式

$$\mu_i=r+\ (\sigma_{im}/\sigma_m^2)\ (\mu_m-r)\tag{32}$$

一种替代表示为

$$ER_i=r+\left[\frac{\mathrm{Cov}\ (R_i,\ R_m)}{\mathrm{Var}\ (R_m)}\right]\cdot\ (ER_m-r)\tag{33}$$

当允许借入或贷出无风险资产时，为了使投资者愿意持有资产 *i*，他们在市场上必须获得预期收益或必要收益

$$ER_i=r+\beta_i\ (ER_m-r)$$

$$\beta_i=\mathrm{Cov}\ (R_i,\ R_m)\ /\mathrm{Var}\ (R_m)\tag{34}$$

我们希望得到对（34）进行重新排列后的表达式。将（23）中的 ER_m-r 代入（34）就得到

$$ER_i=r+\lambda_m\mathrm{Cov}\ (R_i,\ R_m)\tag{35}$$

5.4　β 与系统风险

如果我们将资产 *i* 超过无风险利率的超额收益定义风险溢价，

$$ER_i=r+rp_i\tag{36}$$

那么 CAPM 模型给出了风险溢价的如下表达式

$$rp_i=\beta_i\ (ER_m-r)\ \ =\lambda_m\mathrm{Cov}\ (R_i,\ R_m)\tag{37}$$

CAPM 模型预言，只有资产与市场投资组合收益之间的协方差才对（资产）横截面上的超额收益产生影响。包括红利 - 价格比率、公司规模、收益价格比等在内的其他变量不会对横截面上的期望超额收益产生影响。资产 *i* 的风险变化包含在协方差 Cov（R_i，R_m）的变化之中。严格地说，这个协方差为条件协方差，在任何时点上，代表性投资者都会形成对协方差/β 值的最好看法。

证券市场线

可以对 CAPM 模型进行“重新排列”，可以根据证券市场线（*SML*）来进行表示。假设

$ER_m - r$ 的历史平均值为 8%，无风险利率为 5%，CAPM 就变为

$$ER_i = r + 8\beta_i \tag{38}$$

这是横截面上平均收益 ER_i 与风险 β_i 之间的线性关系（图 8），也是众所周知的证券市场线（*SML*）。如果 CAPM 模型正确，那么所有证券都位于证券市场线上。

根据 CAPM/SML，每种资产的平均月度超额收益（比方说在 60 个月的时期上）$ER_i - r$ 应与资产的 β 成正比，β 越高的证券风险就越大，所获得平均收益也要高一些：

$$\frac{\text{证券 } i \text{ 的超额收益}}{\text{证券 } j \text{ 的超额收益}} = \frac{ER_i - r}{ER_j - r} = \frac{\beta_i}{\beta_j} \tag{39}$$

可以用证券市场线来寻找那些价值被低估或高估的股票。为了看清这一点，考虑一种 β 为 0.5 的证券 *S*（图 8）。你可以通过如下投资组合来复制证券 *S* 的 β：持有 50% 的无风险资产（$\beta = 0$）和 50% 的 β 为 1 的证券（即 $\beta_p = 0.5\ (0)\ + 0.5\ (1)\ = 0.5$）。所构造的这个投资组合应位于 SML 上（*P* 点），它比证券 *S* 有更高的期望收益。于是，*S* 将被卖掉，因为它的实际收益比给定 β 值时的均衡收益低。如果 *S* 被卖掉，它的当前价格就会下降，这就使期望收益增加，从而 *S* 点向 *P* 点移动。（类似地，$\beta_i = 1.2$ 的证券 *T*，可以通过如下方式来进行复制：以无风险利率借入总财富 20% 的资金，与自有资金一起投资于 $\beta_i = 1$ 的证券。）

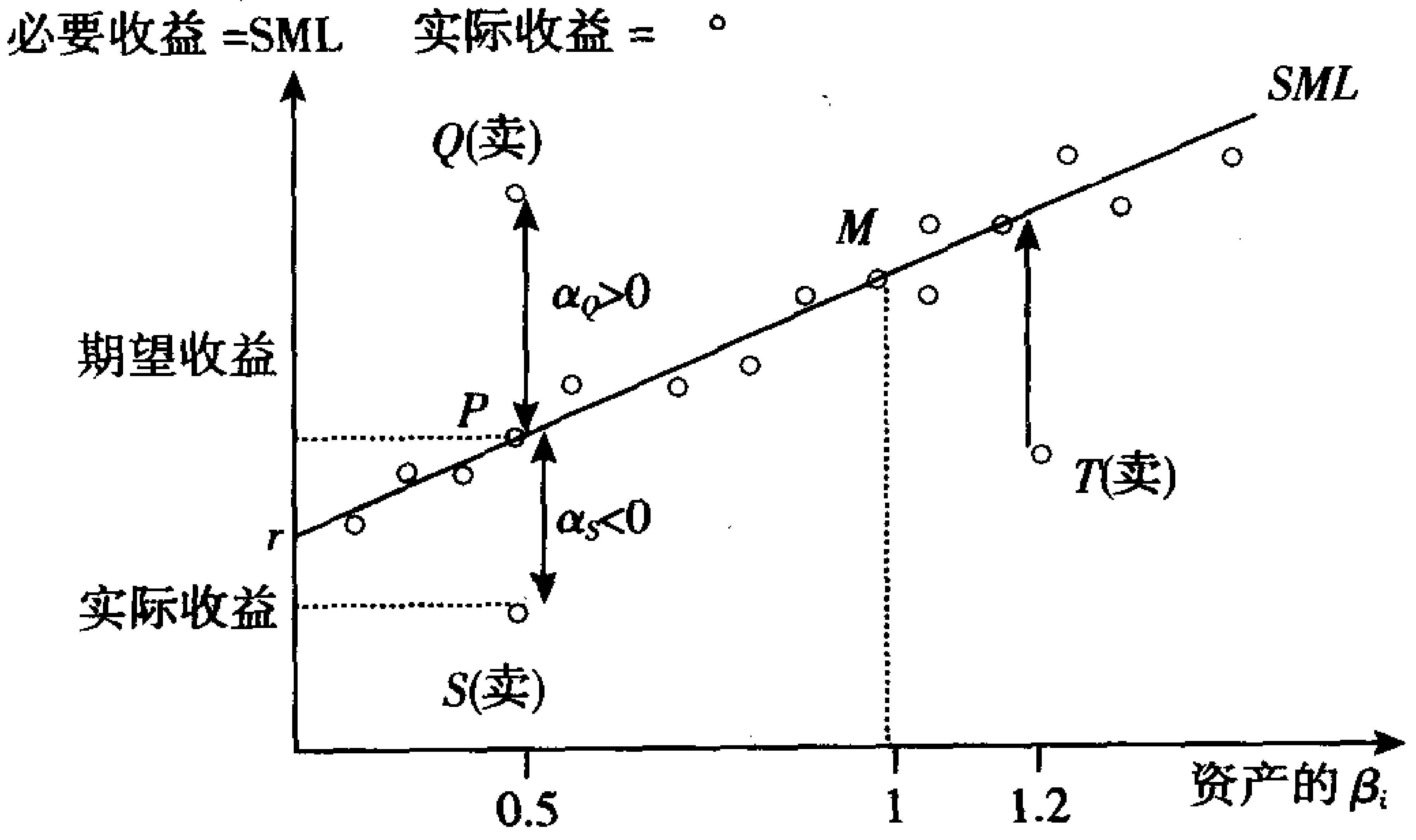


图 8　证券市场线

SML 上（下）证券的"α"为正（负）意味着，这些证券在对"β 风险"进行调整后仍然存在正（负）的"异常收益"。

作为替代，考虑证券 *Q*，它的 β_i 也等于 0.5（图 8），当前的实际收益高于 SML 所表示的平均收益，投资者应该买入证券 *Q*。证券 *Q* 和证券 *S* 在当前时刻的定价都是错误的（即他们没在 SML 上），投机者可能会用卖空 *S* 的资金去购买证券 *Q*。如果这种错误定价被纠正，那么证券 *Q* 的价格将会上涨，因为它在当前时刻有高于平均收益的"异常收益"。所以每个人都会购买证券 *Q*。反之，每个人都会卖空证券 *S*，因为证券 *S* 的市场价格会下降。如果你是第一个发现这种错误定价的人，并且在所有人之前进行交易，那么你就可以从错误定价中获得可观的利润。一个细微之处是，如果市场收益（$R_m - r$）在这一时期中意想不到地上升或下降，那么你仍然可以获得利润。（你能看出为什么吗？提示：注意 *S* 和 *Q* 都有相同的 β。）为了实施这个多头 - 空头策略，人们必须"画出"一组证券的历史平均收益 $\bar{R}_i$（比方说，过去 5 年内的月度平均收益）和 β_i 估计值之间的关系，然后寻找"异常值较大的那些证券"，如证券 *S* 和证券 *Q*。然而，在实践中应记住，这种投资策略是有风险的，因为必须假设 CAPM 模型正确，β_i 的值是正确度量的，任何错误定价都会在一个合理的短期时间内得到纠正。

系统风险

投资组合系统风险的定义是，增加投资组合中证券的数量无法分散风险（这就是众所周

知的不可分散风险、"组合风险"或"市场风险")。即使在充分分散的投资组合里，这种风险也始终存在，因为收益的协方差存在。我们有

$$\sigma_p^2 = \sum_{i=1}^{n}\omega_i^2\sigma_i^2 + \sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n}\omega_i\omega_j\sigma_{ij} \tag{40}$$

为了说明投资组合方差对协方差项的依赖性，考虑一个简单的投资组合：以相同的比例（$x_i = 1/n$）持有所有资产，所有方差和协方差都为常数（即 $\sigma_i^2 = \mathrm{Var}$，$\sigma_{ij} = \mathrm{Cov}$，其中"Var"和"Cov"都是常数）。于是，(40) 变为

$$\sigma_p^2 = \frac{n}{n^2}\mathrm{Var} + \frac{n\ (n-1)}{n^2}\mathrm{Cov} = \frac{1}{n}\mathrm{Var} + \ \left(1-\frac{1}{n}\right)\ \mathrm{Cov} \tag{41}$$

从而，当 $n\to\infty$ 时，每种证券的方差对投资组合方差的影响接近于0，所以投资组合的方差完全由（常数）协方差来决定。于是，个体证券收益的方差就"完全被分散掉"了，但协方差项是无法分散的，协方差项（初略地讲）导致了系统风险，我们用 β 来表示证券的风险。因此，当资产被当作充分分散风险投资组合的一部分持有时，个体证券的异质型风险（σ_i）并不能获得任何收益，因为通过分散化投资的方式可以将这种风险进行分散。只有系统风险——用资产的 β 来表示——才能"获得"期望收益（即"只有 β 风险才有定价"）。

我们可以说明 β 值与它对投资组合风险的贡献之间的如下联系。最优（市场）投资组合的方差为

$$\begin{aligned}\sigma_m^2 &= \omega_1\ (\omega_1\sigma_{11} + \omega_2\sigma_{12} + \omega_3\sigma_{13} + \cdots + \omega_n\sigma_{1n})\\ &\quad + \omega_2\ (\omega_1\sigma_{21} + \omega_2\sigma_{22} + \omega_3\sigma_{23} + \cdots + \omega_n\sigma_{2n})\\ &\quad + \omega_3\ (\omega_1\sigma_{31} + \omega_2\sigma_{32} + \omega_3\sigma_{33} + \cdots + \omega_n\sigma_{3n})\\ &\quad + \cdots\\ &\quad + \omega_n\ (\omega_1\sigma_{n1} + \omega_2\sigma_{n2} + \omega_3\sigma_{n3} + \cdots + \omega_n\sigma_{nn})\end{aligned} \tag{42}$$

我们在这里将 σ_i^2 改写为 σ_{ii}。如果 ω_i 是证券在市场投资组合中的比例，那么我们可以将均衡时的方差记为 σ_m^2。例如，证券2对投资组合方差所作的贡献可以解释为 (42) 中第二行括号中的项，比例 ω_2 就是证券2在投资组合中的"权重"。括号中的项包括证券2和其他证券之间的协方差，也包括证券2自身的方差，每一个协方差 σ_{ij} 都由资产在市场组合中的比例来确定。很容易证明，(42) 中第二行括号中的项就是证券2的收益和市场投资组合收益 R_m 之间的协方差：

$$\begin{aligned}\mathrm{Cov}\ (R_2,\ R_m) &= E\ (R_m-\mu_m)\ (R_2-\mu_2)\\ &= E\sum_{i=1}^{n}\omega_i\ (R_i-\mu_i)\ (R_2-\mu_2)\\ &= \omega_1\sigma_{12} + \omega_2\sigma_{22} + \omega_3\sigma_{32} + \cdots + \omega_n\sigma_{n2}\end{aligned} \tag{43}$$

很容易证明，证券2对投资组合风险的贡献可以由上式给出，因为 $\frac{\partial\sigma_m^2}{\partial\omega_2} = 2\mathrm{Cov}\ (R_2,\ R_m)$。一般而言，市场投资组合的方差可以写为：

$$\sigma_m^2 = \omega_1\mathrm{Cov}\ (R_1,\ R_m)\ + \omega_2\mathrm{Cov}\ (R_2,\ R_m)\ + \cdots + \omega_n\mathrm{Cov}\ (R_n,\ R_m) \tag{44}$$

现在，将 β_i 的定义表达式进行重新排列后就得到

$$\mathrm{Cov}\ (R_i,\ R_m)\ = \beta_i\sigma_m^2 \tag{45}$$

将 (45) 代入 (44)，我们有

$$\sum_{i=1}^{n}\omega_i\beta_i = 1 \tag{46}$$

于是，证券的 β_i 值可以就证券对投资组合风险的增加效应进行度量。当我们将 $\beta_i = 0$ 的证券加入到投资组合中后，这些证券对投资组合方差没有额外的影响。当 $\beta_i < 0$ 时，这种操作却可以降低投资组合的方差。当然，持有证券 i 的数量越大（即 ω_i 的绝对值越大），在其他条件相同的情况下，β_i 对整个投资组合方差的影响就越大。因为 β_i 值较小的资产能大大地降低风险组

合的整体方差，所以人们愿意持有这种资产，尽管它的期望收益相对较低。所有投资者都厌恶风险喜欢期望收益，他们都要在风险和收益之间进行权衡。因此，均衡时，人们愿意持有能降低整体投资组合风险但收益较低的资产。

均衡收益的可预测性

在这一节里，我们将说明 CAPM 模型如何与可变、可预测的收益相一致。如果将 CAPM 模型用于市场投资组合，那么均衡期望（超额）收益为（参见（23）和第 6 章的（27））：

$$E_t R_{m,t+1} - r_t = \lambda E_t\ (\sigma^2_{m,t+1}) \tag{47}$$

加入下标 t 是为了强调这些变量随着时间会发生变化这一事实。由（47）可以看出，如果收益预测误差的条件方差具有时变性，那么均衡超额收益也具有时变性。从理论角度看，无论条件方差是否具有时变性，CAPM 模型都是一个静态模型。由于争论方面的原因，假设如下现象是一个经验事实：股票市场动荡时期或不确定性较大的时期之后，通常出现的是动荡更大的时期。类似地，假设平静时期之后是更为平静的时期。一种说明波动率持久性的简单数学方法是，假设波动率服从 1 阶自回归（AR（1））过程，简单的表示方法是 GARCH（1，1）模型（参见第 2 章和第 29 章的完整讨论）。经济文献通常将条件方差写成 $h_{t+1} \equiv \sigma^2_{t+1}$：

$$\sigma^2_{t+1} = \alpha\sigma^2_t + \beta\varepsilon^2_t \tag{48}$$

这里的 $\varepsilon_{t+1} \equiv R_{t+1} - E_t R_{t+1}$ 是（方程（47）中）零均值收益的条件预测误差。σ^2_{t+1} 在 t 时刻的最佳预测是 $E_t\sigma^2_{t+1} = (\alpha+\beta)\ \sigma^2_t$。CAPM 模型和 ARCH 模型合在一起就意味着

$$E_t R_{m,t+1} - r_t = \lambda\ (\alpha+\beta)\ \sigma^2_t \tag{49}$$

因此，在我们的均衡模型中，期望收益具有时变性，且依赖于 t 时刻的信息 σ^2_t。期望收益随着 σ^2_t 的变化而发生变化的原因是非常直接的。条件方差 σ^2_t 是投资者对下一期市场系统风险 $E_t\sigma^2_{t+1}$ 的最佳估计。均衡时，这种风险可以得到一个较高期望收益回报。

上述模型可以与我们的简单假设——均衡期望收益为常数——进行对比。例如，发现实际收益依赖于 t 时刻或之前的信息 Ω_t（如红利价格比）就可以拒绝后一个模型，这可能是因为，Ω_t 中的变量与省略掉的变量 σ^2_t 相关，在“正确”的预期收益模型（即 CAPM + ARCH）中出现的就是这种情形。上述关于收益可预测性的结论对个体资产的均衡超额收益而言也是成立的。

$$E_t R_{i,t+1} = \lambda \mathrm{Cov}_t\ (R_{i,t+1},\ R_{m,t+1}) \tag{50}$$

$\mathrm{Cov}_t\ (R_{i,t+1},\ R_{m,t+1})$ 为条件协方差。如果根据 t 时刻的信息，协方差是部分可预测的，那么资产 i 的均衡收益就不是常数，而且可以预测。因此，收益可预测的经验证据并不必然意味着投资者是非理性的或投资者忽略了潜在的市场获利机会。

5.5 小结

* 如果投资者对期望收益和方差都有相同的预期，那么不管投资者对风险和收益的偏好怎样，所有“均值方差”投资者都以相同的比例持有风险资产。按照最优比例持有风险资产就构成了市场投资组合。

* 投资者的偏好会进入决策过程的第二个阶段，即在一组风险资产和一种无风险资产之间进行选择。更加厌恶风险的投资者，在风险资产上的财富投资数量也就越小，但持有风险资产的比例与风险厌恶程度低一些的投资者相同。

* CAPM 模型（和 SML）是均值方差投资组合理论的一个结论。CAPM 模型表明，均衡时，单个风险资产的期望超额收益 $ER_i - r$ 都与市场组合的超额收益 $ER_m - r$ 成比例。常数比例就是资产的 β 值，$\beta_i = \mathrm{Cov}\ (R_i,\ R_m)\ /\sigma^2_m$。

* 资产 i 和资产 j 之间的相对期望超额收益由两种资产的 β 值唯一确定。

* CAPM模型并不必然意味着均衡收益是常数。如果条件方差和协方差 $\mathrm{Cov}_t(R_i, R_m)$ 都是时间的函数，那么均衡收益也是时间的函数。另外，根据 t 时刻所获得的信息可以对均衡收益进行预测。如果CAPM模型中用于度量风险的方差、协方差本身是可预测的，那么均衡收益也是可预测的。

第6章

分散化国际投资组合

目的

* 当资产数目较大时，解释计算最优投资组合权重的数学方法，推导 CAPM 模型。

* 考察最优投资组合权重对平均收益、方差、协方差估计值变化的敏感性，考察这个问题的实际解法。

* 在增加外国资产后，说明可以使有效前沿“向左移动”，给出更好的风险-收益权衡关系，考察实践中的应用问题。

* 分析国际投资者是否可以用所持有的外国资产来对货币收益进行套期保值。如果可以，是选择远期、期货还是期权？

这一章分析均值方差组合理论应用到资产配置领域的理论和实践问题。

第一，我们知道，从大的股票集合（如 S&P500 中的股票）中随机地选择股票时，当投资组合所包含的股票约为 25 只时，投资组合的风险（用标准差来度量）接近于最小值。然而，某些风险仍然存在，这就是众所周知的系统风险或不可分散风险。由此产生的问题是，将外国股票（或资产）加入到所选择的投资组合中去是否可以进一步降低系统风险。

第二，是否存在比随机选择股票更好的方法？总的来说，我们感兴趣的是“风险和收益”之间的权衡。一旦我们知道了收益（和预期收益的预测值）的方差和协方差（相关系数），有效前沿就为某组给定的资产给出了这种权衡关系。如果我们将资产的种类数扩大（如同时包括国外资产和国内资产），那么将有效前沿往“左边”大幅移动是可能的，对投资者而言，这种移动能给出更好的风险收益权衡，因为不同国家的资产收益之间存在相对较低的相关性。分散化投资组合能带来好处的一个关键问题是，过去较低的相关性水平在将来是否继续存在，随着时间的演变，波动率和（特别是）平均收益的估计值是否保持不变。

第三，如果我们的投资组合包含国外资产，那么投资者通常感兴趣的是用“本国货币”度量的收益（和风险）。如果投资者没有对这些风险进行套期保值，那么我们需要估计双边汇率下的方差和相关系数；如果投资者对这些风险进行了套期保值，那么我们就可以忽略汇率风险，因为只有“本国货币”的收益和波动率才是重要的。当然，还有一个问题是，使用什么金融工具来进行套期保值，远期、期货还是期权？

第四，如果我们允许投资者以无风险利率借入或贷出资金，我们拥有个体投资者对期望收益和方差的预测值，那么我们就能为这个特殊的投资者确定最优投资组合权重。然后，使用像夏普比这样的绩效指标，我们能对各种国际投资策略（如等权重的世界投资组合）的绩效进行比较。

第五，最优投资组合的权重明显取决于资产预期收益、方差和协方差的估计值，我们需要知道，最优权重对这些重要变量（因为肯定存在度量误差）的敏感性到底有多大，对这些权

重施加约束条件（而不是进行卖空操作）是否有助于问题的解决。

6.1　均值方差模型的数学推导

如果对最优权重不施加任何约束条件（如允许 $\omega_i<0$ 时存在解，也就是允许卖空），那么对单期优化问题而言，为了得到最优投资组合的权重，所需要的数学知识是相当直接的。这是一个存在解析解的标准二次规划问题。如果我们想在一定的约束条件下（如不允许卖空）来求解投资组合问题，那么一般而言不会存在解析解，需要使用数值优化方法，但这些技巧现在都变得很平常了（例如，经常使用 EXCEL 表格和 RATS、GAUSS、Mathematica 等程序）。

这一节说明各种投资组合问题背后的数学推导，投资组合问题中的投资者关心的是投资组合的期望收益和方差。读者应注意的是，尽管所使用的数学知识不超过简单微积分的范围，但根据最大值函数和约束条件，这里所讨论的不同模型还是存在一些细微差别。处理这些差别时要小心谨慎，我们在本节结尾部分对此进行概括。我们将涉及如下主题：

* 为了使投资者位于有效前沿的某个点上，应该如何计算每种风险资产的持有比例 ω_i。这就是著名的 Markowitz 问题。

* 我们说明使用 Markowitz 有效前沿上的任意两点就能画出整个有效前沿。这就是所谓的两基金定理。

* 如何将无风险借贷与 n 种风险资产的选择结合起来以确定市场投资组合的最优持有比例 ω_i。这就是资本市场线与有效前沿的切点。

* 应用两基金定理来说明市场投资组合是两个有效投资组合的某个特殊组合。

* 说明资本资产定价模型（CAPM）是均值方差优化问题的一个直观结论（我们允许投资者以无风险利率进行借贷）。

有效前沿：Markowitz 模型

为了找到有效前沿上的点，投资者必须求解如下带约束条件的最优化问题：

$$\min \frac{1}{2}\sigma_p^2=\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_i\omega_j\sigma_{ij}=\frac{1}{2}(\omega'\Omega\omega) \tag{1}$$

其中，$\omega'=(\omega_1,\omega_2,\cdots,\omega_n)$，$\Omega$ 是 $n\times n$ 的协方差矩阵 $\{\sigma_{ij}\}$。约束条件为：

$$\sum_{i=1}^{n}\omega_i\mathrm{E}R_i=\mathrm{E}R_p \tag{2a}$$

$$\sum_{i=1}^{n}\omega_i=1 \tag{2b}$$

在（1）中，使用1/2仅仅是为了计算方便。注意，投资者拥有 σ_{ij} 和 $\mathrm{E}R_i$ 的估计值，投资者只能将自己的财富进行投资，所以 $\sum_i\omega_i=1$。注意 ω_i 没有任何形式的约束，ω_i 的某些最优值可能小于零（$\omega_i<0$），从而我们允许卖空。这个最优化问题的求解程序是非常标准的，涉及两个约束条件的拉格朗日乘子 λ 和 ψ，

$$\min L=(1/2)\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_i\omega_j\sigma_{ij}-\lambda\left(\sum_{i=1}^{n}\omega_iER_i-ER_p\right)-\psi\left(\sum_{i=1}^{n}\omega_i-1\right) \tag{3}$$

使用矩阵符号很容易对这类问题进行求解（参见附录I），但考虑3个变量情形的代数运算存在教学上的优势：

$$\begin{aligned}L=&(1/2)[\omega_1^2\sigma_{11}+\omega_1\omega_2\sigma_{12}+\omega_1\omega_3\sigma_{13}+\omega_2\omega_1\sigma_{21}\\&+\omega_2^2\sigma_{22}+\omega_2\omega_3\sigma_{23}+\omega_3\omega_1\sigma_{31}+\omega_3\omega_2\sigma_{32}+\omega_3^2\sigma_{33}]\\&-\lambda(\omega_1ER_1+\omega_2ER_2+\omega_3ER_3-ER_p)\\&-\psi(\omega_1+\omega_2+\omega_3-1)\end{aligned} \tag{4}$$

方程（4）关于 ω_i、λ 和 ψ 求一阶偏导就得到一阶条件（FOC）：

$$\frac{\partial L}{\partial \omega_1} = (1/2)\ [2\omega_1\sigma_{11} + \omega_2\sigma_{12} + \omega_3\sigma_{13} + \omega_2\sigma_{21} + \omega_3\sigma_{31}]\ - \lambda ER_2 - \psi = 0 \tag{5a}$$

$$\frac{\partial L}{\partial \omega_2} = (1/2)\ [\omega_1\sigma_{12} + \omega_1\sigma_{21} + 2\omega_2\sigma_{22} + \omega_3\sigma_{23} + \omega_3\sigma_{32}]\ - \lambda ER_2 - \psi = 0 \tag{5b}$$

$$\frac{\partial L}{\partial \omega_3} = (1/2)\ [\omega_1\sigma_{13} + \omega_2\sigma_{23} + \omega_1\sigma_{31} + \omega_2\sigma_{32} + 2\omega_3\sigma_{33}]\ - \lambda ER_3 - \psi = 0 \tag{5c}$$

$$\frac{\partial L}{\partial \lambda} = \sum_{i=1}^{3} \omega_i ER_i - ER_p = 0 \tag{5d}$$

$$\frac{\partial L}{\partial \psi} = \sum_{i=1}^{3} \omega_i - 1 = 0 \tag{5e}$$

最小方差集合所对应的方程

最后两个方程仅仅是约束条件的一个复述。然而，前三个方程的模式是一致的。注意，$\sigma_{ij} = \sigma_{ji}$，上述方程可以扩展到 n 种资产的情形，

$$\sum_{j=1}^{n} \sigma_{ij}\omega_j - \lambda ER_i - \psi = 0 \tag{6a}$$

或者

$$\Omega\omega - \lambda ER - \psi \cdot e = 0 \tag{6b}$$

$$\sum_{i=1}^{n} \omega_i ER_i = ER_p \tag{6c}$$

且

$$\sum_{i=1}^{n} \omega_i = 1 \tag{6d}$$

其中，Ω 是 $n \times n$ 的协方差矩阵，ER 是 $n \times 1$ 向量，e 是 $n \times 1$ 单位向量，ψ 和 λ 是两个数。在方程（6a）中，σ_{ij}和 $ER' =$（ER_1，ER_2，…，ER_n）是已知的；在方程（6c）中，我们将 ER_p 任意地设定为一个固定值。我们有（$n+2$）个线性方程和（$n+2$）个未知数（ω_i、λ 和 ψ）。一旦知道一些基础变量的取值，即这些资产的均值 ER_i（期望收益的预测值）、标准差和协方差 σ_{ij}，使用表格程序（如在 Excel 中），就很容易可以解出这些线性方程，从而得到最小方差点上的最优权重。将所得到的最优权重 ω_i（$i=1, 2, \cdots, n$）代入 $\sigma_p^2 = \omega'\Omega\omega$ 和 $ER_p = \omega' ER$ 就得到有效前沿上的一个点。选择一个新的（任意的）ER_p 值，重复上面的求解过程，直到有足够多的点来画出有效前沿时为止（参见网上的 Excel 电子表格）。

因为这些最优权重完全不受约束，所以可以任意取值，负值表示卖空股票。卖空股票所获得的资金可以投资到其他资产上去，因此，某些资产的 ω_i 值可能大于 1（即投资于这种资产的资金超过了投资者初始时刻财富中这种资产的持有量）。最优权重的唯一约束就是所有 ω_i 的和为 1，这意味着，投资者将所有初始财富都投资于风险资产。显然，上面的计算过程说明了如何构造有效前沿，但这不是“有效”的计算方法。然而，利用方程（6a）～（6d）解的特性就能以更为简单的方式来对问题进行求解，这就是著名的两基金定理。

两基金定理

对 $ER_p^{(1)}$ 和 $ER_p^{(2)}$ 的任意值而言，假设我们找到了 Markowitz 问题一阶条件（FOC）（方程（6a）～（6d））的两个解。

解 1：

$$\omega^{(1)} = \{\omega_1^{(1)}, \omega_2^{(1)}, \cdots, \omega_n^{(1)}\}', \lambda^{(1)}, \psi^{(1)}$$

解 2：

$$\omega^{(2)} = \{\omega_1^{(2)}, \omega_2^{(2)}, \cdots, \omega_n^{(2)}\}', \lambda^{(2)}, \psi^{(2)}$$

这里的 $\omega^{(1)}$和 $\omega^{(2)}$是列向量。于是，$\omega^{(1)}$和 $\omega^{(2)}$的任意线性组合都是 FOC 的解，从而方程（6a）

~（6d）的另一个解是 $\omega^{(q)}=\alpha\omega^{(1)}+(1-\alpha)\omega^{(2)}$，$-\infty<\alpha<\infty$，很容易看出，$\sum_i\omega_i^{(q)}=1$。因为 $\omega^{(1)}$ 和 $\omega^{(2)}$ 两个解都使方程（6a）的左边等于0，所以它们的线性组合也满足（6a）。因为我们已经解出了 $\omega^{(1)}$ 和 $\omega^{(2)}$，所以改变 α 的值就可以画出整个最小方差集合（包括曲线上半部分有效集合上的所有点）。两基金定理意味着，如果投资者能找到两只“共同基金”（即位于有效集合上的证券投资组合，$\omega^{(1)}$ 和 $\omega^{(2)}$ 的资产比例满足 Markowitz 方程），那么将两只共同基金以（α，$1-\alpha$）的比例组合在一起就能构造出均值方差前沿上的任意投资组合。简言之：

两基金定理意味着，为了画出整个有效前沿，我们只需要有效前沿上的两个点（经常称为两只共同基金）就足够了。

投资组合 q 的期望收益和方差为：

$$ER^{(q)}=\sum_{i=1}^{n}\omega_i^{(q)}ER_i=\omega^{(q)\prime}ER \tag{7a}$$

$$[\sigma^{(q)}]^2=\omega^{(q)\prime}\Omega\omega^{(q)}=\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_i^{(q)}\omega_j^{(q)}\sigma_{ij} \tag{7b}$$

因为 $\omega^{(q)}$ 是关于 α 的函数，所以（任意）选定一个 α 值就给出了均值方差有效前沿上的一个点。从纯粹数学角度来看，两基金定理能使 Markowitz 方程（6a）~（6d）的求解程序得以简化。为了针对 ER_p 的所有值来对（6a）~（6d）进行求解，我们只需找到任意两个解就可以了，这两个解的线性组合就导出了所有的解。使问题简化的明显选择是：

（a）　$\lambda=0$，$\psi=1$

（b）　$\psi=0$，$\lambda=1$

在“(a)”中，可以忽略约束条件 $\sum_i\omega_iER_i=ER_p$，因此这就给出了与最小方差组合相对应的 $\omega^{(1)}$。施加（a）和（b）的约束之后，约束条件 $\sum_i\omega_i=1$ 会被违背，但在对解进行尺度变换后，这个约束条件又得以满足。在（6a）或（6b）中，令 $\lambda=0$ 和 $\psi=1$ 就得到了 n 个未知数 $Z_{(i)}^{(1)}$ 的 n 个方程（$i=1, 2, \cdots, n$）。

$$\sum_{j=1}^{n}\sigma_{ij}z_j^{(1)}=1 \quad 或 \quad \Omega z^{(1)}=\mathrm{e} \tag{8a}$$

很容易解出 $Z^{(1)}=\{z_1^{(1)}, z_2^{(1)}, \cdots, z_n^{(1)}\}$，再对 z_i 进行标准化处理后，它们的和等于1

$$\omega_i^{(1)}=\frac{z_i^{(1)}}{\sum_{j=1}^{n}z_j^{(1)}} \tag{8b}$$

于是，向量 $\omega^{(1)}=\{\omega_1^{(1)}, \omega_2^{(1)}, \cdots, \omega_n^{(1)}\}$ 就是有效前沿上的一个最小方差点。第二个（任意）解满足 $\psi=0$ 和 $\lambda=1$，方程（6a）变为

$$\sum_{j=1}^{n}\sigma_{ij}z_j^{(2)}=ER \quad 或 \quad \Omega z^{(2)}=ER \tag{9}$$

这就给出了解 $Z^{(2)}=\{z_1^{(2)}, z_2^{(2)}, \cdots, z_n^{(2)}\}$，$\omega^{(2)}=\{\omega_1^{(2)}, \omega_2^{(2)}, \cdots, \omega_n^{(2)}\}$。在得到解 $\omega^{(1)}$ 和 $\omega^{(2)}$ 以后，我们就能计算出 $\sum_i\omega_iER_i=ER_p$ 和 $\sigma_p=\sum_{i,j}^{n}\omega_i\omega_j\sigma_{ij}$ 在这两个有效前沿点上的值。然后，对任意的 α，计算 $\omega^{(2)}=\alpha\omega^{(1)}+(1-\alpha)\omega^{(2)}$ 的值；改变 α 的值并使用（7）就可以画出（$ER^{(q)}$，$\sigma^{(q)}$）的所有组合。

例子

考虑5种资产的情形（参见网上的 Excel 文档）。假设（使用历史收益数据的样本平均值）已经得到了 σ_{ij} 和 ER 的如下值：

$$\Omega_{5\times5}=\{\sigma_{ij}\}=\begin{bmatrix} 2.2 & 0.9 & -0.3 & 0.65 & -0.42 \\ 0.9 & 1.5 & -0.39 & 0.2 & 0.47 \\ -0.3 & -0.39 & 1.8 & 0.8 & 0.27 \\ 0.65 & 0.2 & 0.8 & 1.5 & -0.5 \\ -0.42 & 0.47 & 0.27 & -0.5 & 1.7 \end{bmatrix} \tag{10a}$$

$$ER' = (ER_1,\ ER_2,\ \cdots,\ ER_5) = [8.5,\ 18.3,\ 12.7,\ 10.8,\ 9.5] \tag{10b}$$

用矩阵符号来表示，$n=5$ 时的方程（8a）可以写为：

$$z^{(1)} = \Omega^{-1} e \tag{11}$$

这里的 e 是 5 个元素都为 1 的 5×1 列向量，因此

$$z^{(1)} = \Omega^{-1} e = \{0.3902,\ 0.2073,\ 0.3091,\ 0.5519,\ 0.7406\} \tag{12a}$$

$$\sum_{i=1}^{5} z_i^{(1)} = 2.1991 \tag{12b}$$

$$\begin{aligned} \omega_i^{(1)} &= z_i^{(1)} / \sum z_i^{(1)} \\ &= \{\omega_1,\ \omega_2,\ \omega_3,\ \omega_4,\ \omega_5\} \\ &= \{0.1774,\ 0.0942,\ 0.1406,\ 0.2510,\ 0.3368\} \end{aligned} \tag{12c}$$

为了得到最小方差点，我们所要求的权重实际上就是 $\omega^{(1)}$（因为在方程（3）中，我们假设 $\lambda=0$）。使用方程（9）就得到第二个解，用矩阵符号表示为：

$$\Omega z^{(2)} = ER \tag{13a}$$

$$z^{(2)} = \Omega^{-1} ER \tag{13b}$$

$$= \{-1.3839,\ 16.1736,\ 10.6532,\ -0.3818,\ 0.3368\}$$

$$\omega^{(2)} = \{-0.0576,\ 0.6730,\ 0.4433,\ -0.0159,\ -0.0428\}$$

当 α 在（$-\infty$，∞）的范围内变化时，使用线性组合 $\omega^{(q)} = \alpha\omega^{(1)} + (1-\alpha)\,\omega^{(2)}$ 来计算 $ER^{(q)}$ 和 $\sigma^{(q)}$ 的值（$ER^{(q)}$、$\sigma^{(q)}$ 都是关于 α 的函数），我们就可以画出整个均值方差集合（从而画出有效前沿）。投资组合方差的无约束最小化问题经常得到组合权重 $\omega_i<0$（即卖空）的解，但有些金融机构是不允许卖空的（如英国的养老基金）。于是，我们需要导出所有 $\omega_i>0$ 的有效前沿。

不允许卖空

在这种情形中，最小化问题与（1）和（2）中的最小化问题相同，但增加了一个约束条件，即 $\omega_i \geqslant 0$（$i=1,\ 2,\ \cdots,\ n$）。现在，我们不能使用两基金定理来简化求解过程，事实上需要使用一种特殊形式的数学规划即二次规划来进行求解，它的一阶条件就是著名的库恩—塔克（Kuhn - Tucker）条件。有些读者可能熟悉这类求解过程。通常出现的情形是，在允许卖空时，几乎所有 ω_i 都是非零的（即要么为正，要么为负），但在不允许卖空时，最优权重 ω_i 中的许多值都为 0，因此，投资者只是持有总资产的一个部分。

借贷情形：市场投资组合

当我们允许投资者以无风险利率进行借或贷来对 n 种风险证券进行投资时，最优解就是市场投资组合（如果所有投资者对期望收益、方差、协方差都持有相同的看法）。变换线和有效前沿的切点确定了最优投资比例（允许卖空）。从数学角度来看，为了得到市场投资组合，我们选择权重 ω_i 来使夏普比 θ 达到最大：

$$\max\theta = \frac{ER_p - r}{\sigma_p} \tag{14}$$

约束条件为：

$$ER_p = \sum \omega_i ER_i \tag{15a}$$

$$\sum \omega_i = 1 \tag{15b}$$

$$\begin{aligned} \sigma_p &= \Big(\sum_{i=1}^{n} \omega_i^2 \sigma_i^2 + \sum_{i\neq j}\sum \omega_i\omega_j\sigma_{ij}\Big)^{1/2} \\ &= \Big(\sum_{i=1}^{n}\sum_{j=1}^{n} \omega_i\omega_j\sigma_{ij}\Big) \end{aligned} \tag{15c}$$

ER_i 是资产 i 的期望收益，ER_p 和 σ_p 分别是投资组合的期望收益和标准差。约束条件（15b）

可以直接嵌入到（14）的最大值函数中去，我们将 θ 写成：

$$\theta = \frac{\sum \omega_i (ER_i - r)}{\sigma_p} \tag{16}$$

可以证明，一阶条件具有如下形式（参见附录 II）：

$$z_1\sigma_{11} + z_2\sigma_{12} + \cdots + z_n\sigma_{1n} = ER_1 - r$$

$$z_1\sigma_{12} + z_2\sigma_{22} + \cdots + z_n\sigma_{2n} = ER_2 - r$$

……

$$z_1\sigma_{1n} + z_2\sigma_{2n} + \cdots + z_n\sigma_{1n} = ER_n - r \tag{17}$$

其中，$z_i = \eta\omega_i$，η 为常数。常数 η 对解不会产生任何影响，因为如果 ω_i 是（16）的一个解，那么 $\eta\omega_i$ 也是（16）的解，η 可以同时从分子和分母中约去。求出方程（17）的解 z_i 后，根据下式就可以确定 ω_i 的最优值：

$$\sum_{i=1}^{n} \omega_i = 1 = \eta^{-1} \sum_{i=1}^{n} z_i \tag{18}$$

于是

$$\eta = \sum_{i=1}^{n} z_i$$

$$\omega_i = \frac{z_i}{\eta} = \frac{z_i}{\sum_{i=1}^{n} z_i}$$

因为 ER_i、r、σ_i^2 和 σ_{ij} 都是已知的，所以方程（17）是由 n 个方程所组成的方程组，由此可以解出 n 个未知数 z_i 到 z_n。方程（17）可以写为（$k = 1, 2, \cdots, n$）

$$\sum_{i=1}^{n} \sigma_{ki} z_i = (ER_k - r) \quad \text{或} \quad \Omega z = ER - r \tag{19}$$

其中，$Z = (z_1, z_2, \cdots, z_n)$，$ER - r = (ER_1 - r, ER_2 - r, \cdots, ER_n - r)'$，$\Omega$ 是 $n \times n$ 协方差矩阵。于是，

$$z^* = \Omega^{-1} (ER - r) \tag{20}$$

很容易得出上述解。我们再一次对最优权重的取值没有施加任何约束条件，如果禁止卖空，那么 $\omega_i \geqslant 0$（对所有的 i），从而优化问题涉及库恩—塔克条件的应用。在实践中，我们能得到各种类型的优化程序，如在通常的工作表程序（如 Excel）里添加一些程序。当资产数目增加时，或当 ω_i 的约束条件变得越来越复杂时，工程师就要使用更有弹性（速度更快）的软件包，如 GAUSS、Excel - Visual Basic、C + + 等等。

在有 n 种资产的一般情形中（加上无风险资产），我们在（19）中不得不求解 n 个方程。一旦我们知道协方差矩阵 Ω 和期望收益的预测值 ER_i，求解过程是相当直接的（在 Excel 中）。优化问题的解为

$$z^* = \{-3.33, 15.14, 9.11, -3.14, -4.73\}$$

$$\omega^* = \{-0.26, 1.16, 0.70, -0.24, -0.36\}$$

前面的 Markowitz 的均值方差问题与市场投资组合的权重之间存在一个很好的联系。市场投资组合所对应的（17）中的一组一阶条件，与 Markowitz 两基金问题（8a）和（9）的一阶条件非常相似。当解为 $z^{(1)}$ 和 $z^{(2)}$ 时，Markowitz 两基金问题的一阶条件可以写为：

$$\sum_{j=1}^{n} \sigma_{ij} z_j^{(1)} = 1 \quad \text{或} \quad \Omega z^{(1)} = e \tag{8a}$$

$$\sum_{j=1}^{n} \sigma_{ij} z_j^{(2)} = \mathrm{ER}_k \quad \text{或} \quad \Omega z^{(2)} = \mathrm{ER} \tag{9}$$

这里的 e 为单位向量。方程（19）的右边为 $\mathrm{ER}_k - r$，左边与（8a）和（9）相同。于是，如果 $z^{(1)}$ 和 $z^{(2)}$ 是 Markowitz 有效前沿一阶条件的两个解，那么 $z^* = z^{(2)} - rz^{(1)}$ 是（19）中市场投资组合问题的一个解。这并不令人感到奇怪。我们在前面发现，Markowitz 解的任何线性组合也是

有效前沿上的点。但是，市场投资组合显然位于有效前沿上，所以它也是 $z^{(1)}$ 和 $z^{(2)}$ 的一个线性组合。我们已经计算出 $z^{(1)}$ 和 $z^{(2)}$，因此当 $r=5\%$ 时，市场投资组合问题的解为：

$$z^{*}=z^{(2)}-rz^{(1)}=\{-3.33,\ 15.14,\ 9.11,\ -3.14,\ -4.73\} \tag{21}$$

由此可以得到市场投资组合的权重为

$$\omega^{*}=\{-0.26,\ 1.16,\ 0.70,\ -0.24,\ -0.36\}$$

当然，后者与（20）中直接运用 $z^{*}=\Omega^{-1}\ (ER-r)$ 所得到的解相同。（20）中的解 z^{*} 是两基金定理的另外一个结果。

注意，如果所有投资者对 ER_i 和 σ_{ij} 都有相同的预期，那么（21）就是市场投资组合。如果投资者对上述变量有自己的预期，那么 z^{*} 给出的仍然是最优权重，但只是单个投资者的最优权重。不同投资者的最优投资组合权重也不同。人们也可以引入其他约束条件，如不允许卖空（即 $\omega_i \geqslant 0$），但我们不能再使用上述简化求解过程，所得到的最优权重也不再是“市场投资组合”中的权重。上面的所有计算都要求输入一些变量，如期望收益（当前无风险利率为5%）、方差、协方差（相关系数），在上面的例子中我们假定这些输入变量都为常数。在实践中，当我们对这些变量的预期发生改变后，我们需要在适当的时期中（如一天、一个月或一年）来估计这些变量，重新计算最优的投资组合比例。

Britten - Jones（1999）给出了一个使用简单回归就能得到最优（无约束）均值方差权重的简洁（事实上很直观的）方法。简单地将因变量 $\boldsymbol{Y}$ 取为元素全为 1 的 $T\times 1$ 列向量，自变量 $\boldsymbol{X}$ 是由 k 种资产的超额收益所组成的 $T\times k$ 矩阵，回归方程没有截距项：

$$\boldsymbol{Y}=\boldsymbol{X}\beta+\mu \tag{22}$$

最优权重为 $w=\boldsymbol{b}/\sum_i^k b_i$，$\boldsymbol{b}$ 为 OLS 估计值。这个程序也能得到 $\boldsymbol{b}$ 的标准误差，所以我们可以对与这些最优权重相关的假设进行检验。特别地，Britten - Jones 考虑了一个最优（没有进行套期保值）分散化国际投资组合问题（从美国投资者的角度看），然后检验非美国资产的权重 ω_i 在统计上是否同时为零。他们发现，美国投资者不可能通过国际投资组合来分散风险。这实际上是对如下结论的一种替代性证明：最优权重对平均收益的假设非常敏感，平均收益的度量却存在很大的不确定性。也需要注意的是，“在本国进行投资”的结论只是对美国投资者进行检验，只有波动率和相关关系数为常数（在总体中）时，这些分析结论才会成立。但他们发现，（对美国投资者而言）最优权重具有时变性（例如，当样本数据样本被分成两个子样本时，德国马克的变化是从 29% 的空方头寸转变为 69% 的多方头寸）。最后，如果事先对权重进行限制（如禁止卖空），那么这些分析结论也不成立。

资本资产定价模型

所有投资者都持有市场投资组合，市场投资组合中所有资产都有一个特定的期望收益。这些期望收益之间有着什么样的联系呢？对任意一个特殊的期望收益而言，哪些因素起着决定性作用呢？CAPM 模型给出了市场投资组合与资产期望收益之间的联系。CAPM 模型假设所有投资者都有相同的预期，都通过（14）对 θ 进行最大化来确定风险资产的持有比例（即在资本市场线和有效前沿的切点上）。当持有所有资产时，任意资产 k 的均衡期望收益为

$$ER_k=r+\beta_k\ (ER_m-r) \tag{23}$$

其中，$\beta_k=\mathrm{Cov}\ (R_k,\ R_m)\ /\sigma_m^2$。上述关系隐含在市场投资组合的一阶条件中。选取（17）中的第 k 个方程，我们有

$$ER_k-r=\eta\ [\omega_1\sigma_{1k}+\cdots+\omega_n\sigma_{nk}] \tag{24}$$

方括号中的项就是 $\mathrm{Cov}\ (R_k,\ R_m)$。因为在最优的 ω_i 值处，我们有

$$\begin{aligned}\mathrm{Cov}\ (R_k,\ R_m)\ &=\mathrm{Cov}\Big[R_k,\ \Big(\sum_{j=1}^{n}\omega_j R_j\Big)\Big]\\&=\sum_{j=1}^{n}\omega_j\mathrm{Cov}\ (R_j,\ R_k)\end{aligned}$$

$$= \sum_{j=1}^{n} \omega_j \sigma_{jk} \qquad (25)$$

因此，方程（13）可以写为

$$ER_k - r = \eta \mathrm{Cov}(R_k, R_m) \qquad (26)$$

因为方程（24）对所有资产都成立，所以它必然对市场投资组合也成立，于是

$$ER_m - r = \eta \sigma_m^2 \qquad (27)$$

将（27）中的η代入（26），我们就得到了 CAPM 模型

$$ER_k - r = [(ER_m - r)/\sigma_m^2] \cdot \mathrm{Cov}(R_k, R_m) = \beta_k (ER_m - r) \qquad (28)$$

任意资产的期望收益都由资产的β值和超额市场收益所决定。在一个分散化投资组合中，任意两种资产的相对风险大小由它们β值的相对大小来决定。

6.2 国际分散化

分散化投资组合的优势取决于“新资产”的寻找，新资产与投资组合中既有资产之间的相关性很“低”。在国际分散化的早期研究中，Solnik（1974）集中关注的是国际分散化可能的风险降低优势（忽视投资组合中的期望收益）。他们提出了如下问题：

* 为了确保一个合理的分散化水平，必须持有多少国内证券？

* 国际分散化降低了风险吗？

使用$\sigma_p^2 = \omega' \Omega \omega$，对每种资产输入相同的权重（$\omega_i = 1/n$），根据简单的历史样本（算术）平均值（使用1966—1971年的每周收益数据）来估计σ_{ij}。分析步骤为：

（1）从一个大的股票集合中随机地选择股票来产生投资组合。

（2）构造由1只股票所组成的投资组合，重复操作m次，计算“1支股票的投资组合”的平均标准差$\sigma_p^{(1)}$（即$\sigma_p^{(1)}$是m个“只有1种资产的投资组合”的平均标准差）。这种平均方法可以避免“异常值”主导计算结果。

（3）对$n=2$，$n=3$，…，重复步骤（2），根据投资组合的“规模”来获得$\sigma_p^{(n)}$的平均值序列。

（4）使用$\sigma_p^{(1)}$来对每一个“基于规模”估计的标准差进行尺度变换：

$$V_p^{(n)} = [\sigma_p^{(n)} / \sigma_p^{(1)}]^2$$

画出$V_p^{(n)}$关于n——投资组合中证券的数目——的图形。（注意，$V_p^{(1)} = 100\%$。）

（5）对不同的国家重复步骤（1）～（4）。

Solnik的重要结论是，在任何国家里随机选择20种“国内证券”就可以将系统（市场）风险降到最低水平。（今天，随着证券之间相关性的增强，要将投资组合的系统风险降到最低水平，美国大约需要30种国内股票，国际投资组合大约需要40种股票。）例如，Solnik发现，对美国和德国而言，$V_p^{(n)}$的最小值分别为27%和44%，这表示德国的市场风险水平更高。

在降低风险方面，国际分散化到底有怎样的效果呢？Solnik假设可以对外币收益进行完全套期保值（远期交易成本为0），从而国外资产的外币收益都是通过远期市场上的售卖来获得的（或借入外币购买股票）。注意，在实践中，上述假设并不能保证你进行了完全套期保值，因为你不能确切地知道你的股票（比方说，在一个月内）究竟值多少钱。对投资德国全股指数DAX的美国居民而言（我们将美国视为“国内”），美元套期保值收益为：

$$\text{套期保值收益} = R_{\mathrm{DAX}} + \text{欧元的远期升水}$$

$$R_{\mathrm{us}}^{h} = R_{\mathrm{DAX}} + (F - S)/S$$

对于美国投资者而言，未套期保值的德国证券投资组合所提供的美元收益为：

$$\text{未套期保值的收益} = R_{\mathrm{DAX}} + \text{欧元升值}$$

$$R_{\mathrm{us}}^{u} = R_{\mathrm{DAX}} + R_S$$

其中，R_{DAX} = DAX 的收益（比例）；R_S = 美元/欧元的收益（即美元/欧元即期利率的比例变化）；$(F-S)/S$ = 欧元的远期升水（汇率用美元/欧元来表示）。

Solnik（1974）选取了九个国家（股票指数），从这些国家中随机地选取股票，来构造基于不同规模的国际投资组合。对没有进行套期保值的投资组合而言，“收益”的标准差和相关系数可以用无套期保值的收益来导出。对套期保值收益而言，假设相对于 R_{DAX} 而言远期升水较小，我们将其设为0。于是，未套期保值的收益等于“本币”收益。然后，Solnik 对套期保值组合和无套期保值组合分别计算了统计指标 $V_p^{(n)}$。

如果相对于单个国家股票收益之间的相关系数而言不同“本币”收益之间的相关性较低，那么套期保值分散化国际投资策略可以降低投资组合的风险 σ_p。Solnik 发现，对美国投资者而言，分散化国际投资组合的统计指标 $V_p^{(n)}$ 下降到 11% 左右（无论是套期保值组合，还是无套期保值组合），这大大低于（美国）国内投资组合的 $V_p^{(n)}$ 值——约为 27%。无套期保值组合和套期保值组合的结果非常接近，这是因为在一个良好的分散化投资组合中，对无套期保值投资组合而言，一组双边汇率（相对于美元而言）运动可以相互抵消。另外，相对于美元的绝大多数汇率的变化不会达到 1974 年的值（当 Solnik 的研究结束时），因为布雷顿森林体系下的半固定汇率制一直存在直到 1973 年。Solnik 的研究在这个领域具有开创性，但也存在明显的局限，即：

* σ_{ij} 是用历史平均值来进行计算的。
* 没有考虑投资组合的预期收益。
* 假设可以对外币收益进行完全套期保值。

对外币收益进行套期保值

直到现在，我们在很大程度上忽略了套期保值问题。考虑一个同时持有国内和国外资产的投资者在远期市场上是否总可以对未来的外币收益进行完全套期保值。答案取决于国内和国外股票市场收益之间的相关性及即期货币之间的相关性。未实行套期保值的头寸是风险（即即期汇率的方差）的另外一个来源。但是，它也可能导致即期汇率之间、国内股票市场、国外股票市场、或国内外股票市场上的低（或负）相关性。这些相关性可能抵消即期汇率的方差，从而能降低整个未套期保值组合的方差。

由上述研究可以明显看出，在对各种投资组合的配置决策进行评价时，有很多潜在的因素需要考虑。这些因素包括：

* 研究对象应包括多少个国家，每个国家应选取多少种资产（如，仅有股票，或仅有债券，或两者都有，或者更为宽泛的资产集合——诸如产权等）

* 在对风险和收益进行度量时，以哪个国家的货币作为记价货币（即本国货币）。本国货币的不同选择或者收益度量时期的不同选择（一个月或一年）对结论不会产生任何影响。

* 我们必须同时考虑期望收益和风险，给出投资组合绩效的测度指标（如夏普比、Treynor 指数）。

* 在实践中，如果使用某些简单的分散化国际投资组合方法（例如，在可供选择的国家之间进行等权重配置，或者根据已有的权重，如根据摩根斯坦利国际资本全球指数（Morgan Stanley's Capital International World Index）来进行配置），那么从中得到好处是可能的。

* 是否对将来的外币收益进行套期保值，使用什么样的金融工具来进行套期保值，远期、期货还是期权。

* 汇率变化的各种预测值（如历史算术平均，随机游走模型，将远期汇率视作下一期的即期汇率）导致了未实行套期保值的不同最优投资组合权重，不同预测值的方差和协方差（相关系数）也不同。

* 如果各种预测方法都有很大的预测误差，那么由不同预测方法所得到的不同最优投资组合权重在统计上不会显著不同。

在这一章里，我们只是简单地涉及这些问题，但 Eun 和 Resnick（1997）结论是值得注意的，因为他们考虑了许多关键问题。他们将美国视为“本国”，使用六个“外国”（加拿大、法国、德国、日本、瑞士和英国）月度股票收益数据，采用简单的历史算术平均方法来对收益的方差和协方差进行度量，因为最优投资组合权重对期望收益的变化相当敏感，所以他们考虑了多种预测方法。他们使用样本内数据来计算这些变量，然后计算最优投资组合权重，接着对“样本期”12个月的事前投资组合收益和标准差进行比较（使用“自助法”（bootstrapping method））。他们将夏普比作为一个测度指标来对选择最优投资组合权重的各种方法所得到的结果进行比较。更为宽泛地讲，他们发现，对美国投资者而言：

* 对未实行套期保值的投资组合而言，分散化国际投资组合的结果要占优于（根据夏普比或随机占优分析）仅投资于美国股票的结果。这个结论与国际投资组合是否包括等权重投资组合无关，与投资组合的权重是否就是摩根斯坦利国际资本全球指数中的权重无关，也与投资组合的权重是否就是均值方差分析所得到的“最优权重”无关。然而，至少对美国投资者而言，未实行套期保值的分散化国际投资策略的收益仅仅在边际上比国内投资要好。（显然，即使未实行套期保值的分散化国际投资策略比纯粹的国内投资策略有着实质性的改进，这个结论对“小国”投资者而言仍然不成立。）

* 在考虑套期保值策略时，使用远期合约进行完全套期保值的结果总比未实行套期保值的结果好（例如，无论使用的是否是“最优”投资组合权重，是否是等权重投资组合，是否是摩根斯坦利国际资本全球指数中的权重，这个结论都成立）。另外，使用远期市场通常比对外币的未来收益实行套期保值的效果好。因此，对美国投资者而言，存在合理的强有力证据，如果投资者进行分散化国际投资，那么他会以某种方式对外币收益的完全套期保值进行支付。

* 人们也发现，如果假设下个月的即期汇率等于本月的即期汇率（即随机游走），那么即期汇率是比当期远期汇率更好的预测指标。当然，这与远期汇率无偏性检验（FRU）中所发现的“糟糕”结果相吻合。在远期汇率的无偏性检验中，即期汇率的变化对滞后远期升水的回归所给出的负回归系数保持不变（参见第25章）。

国内偏好问题

上述证据似乎表明，即使对美国居民而言，进行分散化国际投资也可以获得收益。因此，在实践中就存在一个谜，为什么分散化国际投资是如此的少，尤其是对美国投资者而言。这就是“国内偏好之谜”。

投资组合理论强调，当国内和国外证券收益之间的相关性比纯粹的国内证券之间的相关性低时，或者汇率的运动能降低这些相关性时，分散化国际投资可能获得收益。据估计，美国居民进行国际化投资（在G7国投资）比纯粹的国内投资有10%～50%福利改善。然而，与此相矛盾的是，美国居民90%以上的投资都是“国内”证券。对其他国家的投资者而言结论同样成立，只不过“国内”投资的比重低于90%。例如，英国养老基金目前的国内外投资比例大约为70∶30。这个问题是如此的普遍以至于它以“国内偏好问题”而出名。进行资本控制的风险和高资本税收似乎都不能解释：持有国外股票明显可以获得收益，而投资者却没有利用这样的机会。

“国内偏好”的一个可能原因是，缺乏国外公司绩效方面的详细信息（如小的“外国”公司）。国外投资有时存在法律上的限制。例如，英国养老基金目前不得不满足最低资金要求（MFR），这是对基金偿还债务能力的一种检验，它鼓励投资者投资国内的政府债券，在某种程度上也鼓励投资英国的股票。“国内偏好”的另外一个原因是，使用历史数据“已经证明的收益”并不必然意味着未来（存在不确定性）一定能获得收益。

投资者不进行分散化国际投资的另外一个原因是，他们希望用投资的收益来消费国内商品和服务，当汇率发生波动时，国外投资收益的不确定性就会增加。然而，如果购买力平价理论（PPP）成立（参见第24章），那么国外投资的实际收益与国内居民的实际收益相同，因此

不存在任何外汇汇率风险。不幸的是，购买力平价理论只是在5～10年这样的长期内才成立，短期内实际汇率的变化可能非常的大（如，1～5年上的汇率变化为±25%）。这会引入不确定性，因为当汇率稍微有利时，你就希望将国外投资转变成现金，这可能是“国内偏好”的一个原因。当信息成本降低时（例如，预期外国公司的信息会有所增加，通过网络来获得时时股票价格和即时交易方面的信息），始终可能的情形是，国内偏好问题会有所削弱，投资者可能越来越愿意进行分散化国际投资。

另一方面，有些学者认为，当我们考虑均值方差优化问题中输入变量的测度存在不确定性这个事实后，“国内偏好”问题就不是很牢靠。尽管美国投资者股票中的投资比例与标准普尔500中的比例相同，他还是不可能准确地以均值方差最优比例来持有股票，然而，他与S&P500指数相联系的投资组合可能会落入这个最优头寸95%的置信区间内。在现实世界中，如果我们需要对均值方差优化问题中的输入变量进行“真实”预测（即对期望收益和方差协方差矩阵进行预测），那么这个观点会变得越来越有力。在这种情形中，与标准普尔500指数相联系的投资组合可能比最优的均值方差组合表现得更好。

6.3 实践中的均值方差优化问题

可能出现的情形是，投资基金的绝大部分资金都没有在均值方差优化问题的基础上进行投资。一般而言，投资银行的投资政策委员决定在不同国家进行资产配置时，他们以极其不正规的方式使用各种准则，如政治风险、商业风险、经济周期的状态等。本书所赞成的均值方差优化问题在“纯粹应用”中存在哪些问题？在实践中应用均值方差优化方法存在哪些困难？显然，这种方法只是用方差协方差矩阵来度量风险，但没有涉及其他形式的风险（如政治风险），这种方法（最简单的形式）也只对一个特殊的时期（如，一个月或一年等）有效。然而，这种方法没有获得广泛应用的一个主要原因是，从某种意义上讲它是一个“黑盒子”，这种方法所得到的结果也可能存在较大的估计误差。

估计误差

我们考虑可能的估计误差。如果（连续复利）收益服从独立同分布的正态分布（*niid*），总体的均值μ和方差σ^2都为常数，那么均值和标准差的（无偏）估计值为：

$$\bar{R}=\frac{\sum_{i=1}^{n}R_i}{n} \tag{29a}$$

$$\hat{\sigma}=\sqrt{\frac{\sum_{i=1}^{n}(R_i-\bar{R})^2}{n-1}} \tag{29b}$$

对任意特殊时期上任意频率的数据（如每天、每周、每月）而言，可以用上述公式来计算期望收益和标准差。（由于使用了重叠数据，所以我们忽略一些复杂问题。）例如，使用月度数据可能会发现，每年的$\bar{R}$为1%、每年的$\hat{\sigma}$为4%。可以证明，$\bar{R}$估计值的标准差为：

$$\text{stdv}\ (\bar{R})\ =\frac{\sigma}{\sqrt{n}} \tag{30}$$

在$\sigma=4\%$的条件下，假设我们希望获得1%的总体均值收益的估计值，且估计精度为±0.1%，因此我们需要$n=4^2/(0.1)^2=1600$个月度观测值，即133年的月度数据。显然，使用“合适的样本”，比方说60个月（即5年）的观测值，来获得期望收益的估计值，估计精度将是非常糟糕的。例如，当$n=60$时，stdv（$\bar{R}$）$=\sigma/\sqrt{n}=0.52\%$，估计误差超过了期望收益$\bar{R}$估计值（1%）的一半。还有一个问题是，如果总体的均值不为常数，那么即使使用大量的历史数

据也不能获得精确估计，因为从估计开始的时刻起，数据就不是可变总体均值的准确表示。于是，分析师倾向于使用其他方法来估计期望收益。他们可能会用证券市场线和资产的 β 值来对期望收益进行估计，也可能使用 APT 或更为一般的回归模型来估计期望收益——回归模型假设期望收益取决于一组“基础变量”（如红利收益比或价格收益比、收益率差等）。他们也可能将这些预测结果与公司或子公司报告所提供的辅助信息结合起来分析问题。

当样本标准差为 $\hat{\sigma}=4\%$ 时，估计精度为多少呢？$\hat{\sigma}$（对正态分布收益而言）的标准差为：

$$\text{stdv}\ (\hat{\sigma})\ =\frac{\sqrt{2}\sigma^2}{\sqrt{n-1}} \tag{31}$$

假设使用5年（$n=60$）的数据来对 $\hat{\sigma}=4\%$ 进行估计。根据上述方程，我们得到 stdv（$\hat{\sigma}$）$=\dfrac{\sqrt{2}\times 4^2}{\sqrt{60-1}}=0.38\%$。与每年0.38%（估计值的95%）的标准差相比，$\hat{\sigma}=4\%$ 的估计精度要好一些，使用历史数据对方差（和协方差）估计的误差比对期望收益估计的误差要小。

可以这样认为，保持数据“长度”不变，比方说5年，增加数据的频率，比方说月或天，可以提高对期望收益的估计精度。然而，这种做法事实上是没有任何帮助的——对一个固定的数据集合而言，你不可能无中生有。这是因为，如果（比方说）月度收益在统计上独立，那么可以证明，每日期望收益和标准差可以分别由 $\mu_d=T\mu_m$ 和 $\sigma_d=\sqrt{T}\sigma_m$ 来给出，T 是收益所考察的时间长度在一个月中的比例（这里的 T 近似等于1/30）。于是

$$\sigma_d/\mu_d=T\sigma_m/\sqrt{T}\mu_m=\sqrt{30}\sigma_m/\mu_m$$

这意味着，相对于每日均值的每日标准差大约是相应月度值（即 σ_m/μ_m）的5.5倍。因此，随着时期的缩短相对误差会增加。

黑盒子？

均值方差投资组合分析可能是一个“黑盒子”，因为最优权重 ω_i^* 只是根据最大化程序所“选择出来”的权重，很难用它（特别是资产种类数目很大的时候）来进行易于处理、易于理解的敏感性分析。估计误差的产生是因为，输入变量——期望收益（ER_i）和方差协方差矩阵 $\{\sigma_{ij}\}$ 中的元素——可能是对未来实际发生情况的一个糟糕预测。“优化者”显然会高估（低估）那些有较高（低）期望收益预测值、负（正）协方差和小（大）方差的证券。一般而言，期望收益的预测偏差是误差的主要来源；与此相对照的是，对 σ_{ij} 的预测要合理一些。

一般而言，人们使用过去收益的历史平均值（如，在给定的“数据窗口”下的样本平均收益）来度量期望收益。可以对这些方法进行改进，如使用更为复杂的多元递推回归模型、时变参数模型、纯粹的时间序列模型（如 ARIMA 模型和随机趋势模型）、贝叶斯估计和最近的基于神经网络的预测。在指数加权移动平均（EWMA）模型或者简单的 ARCH 和 GARCH 模型（参见第29章）基础上可以对方差和协方差进行预测。从本质来看，这些方法都假设方差（或协方差）是历史收益平方的加权平均。当然，这些方法所涉及的计算成本会增加，更为重要的是，那些心存疑虑和缺乏专业知识的管理高层解释结果的成本会增加。然而，我们几乎没有证据能说明，在均值方差优化问题中，这些更为复杂的方法到底能将“估计误差”降低多少。Simmons（1999）给出了简单而又能说问题的敏感性分析。他用和方差协方差的历史样本平均值来计算有效前沿（禁止卖空）——投资组合中的资产由美国股票、美国债券、美国货币市场资产、欧洲股票和太平洋证券等组成）——上的最优权重（对美国投资者而言），然后使用 EWMA 方差协方差预测值来重复上述操作，结果发现，最优权重有着实质性的变化，这说明均值方差分析似乎对输入变量的微小变化是极度敏感的。

有什么证据能说明在均值方差优化问题中使用 R_i 和 σ_{ij} 的“历史”样本平均值确实会导致“估计误差”呢（如 Jobson 和 Korkie（1980），Frost 和 Savarino（1988））？正如我们即将看到的那样，一般而言，这些研究都建议，最好的策略就是对任何单一证券的权重进行限制，使权

重取相对较小的值，取值范围是投资组合价值的2% ~5%，并且不允许卖空。

蒙特卡罗模拟技巧可以用来对均值方差优化问题中所隐含的“估计误差”进行度量，这种方法允许使用所生成的资产收益和方差协方差矩阵的“重复样本”。在拟合数据的每一轮计算中，我们都可以计算出最优投资组合的收益 R_p、标准差 σ_p 和夏普比（$R_p - r$）/σ_p 的估计值。然后将这些拟合结果与已知分布中所给出的“真实”值进行比较。整个程序包含以下几个步骤，其中，n 为投资组合的资产数目（如20），m 为蒙特卡罗模拟的次数，q 是在计算平均收益、方差、协方差过程中所使用的样本数据长度。

（1）假设收益服从均值为 μ_i、方差协方差矩阵为 $\Sigma = \{\sigma_{ij}\}$ 的多维正态分布，在两种资产的情形中，μ_1，μ_2，σ_1，σ_2 和 σ_{12} 的“真实值”用 q（=60）个数据点（比方说）的历史样本平均值来进行计算。我们由此假设这些值都是已知常数。因此，我们就知道了真实的“总体参数值” μ_i 和 Σ，我们能计算出使夏普比 $S = (\mathrm{E}R_p^* - r)/\sigma_p^*$ 达到最大的最优权重 ω_i^*、最优投资组合的收益 R_p^* 和 σ_p^* 标准差。

（2）资产收益由多维正态分布生成，多维正态分布包含了资产收益之间的相关性结构：$R_i = \mu_i + \varepsilon_i$。$\varepsilon_i$ 从多维正态分布中进行抽取，方差协方差矩阵 Σ 已知，计算方法同上，由此产生了 q（=60）个拟合收益（R_i 为第 i 种资产的收益，$i = 1, 2, \cdots, n$）。这是蒙特卡罗模拟的第一轮计算（即 $m = 1$））。

（3）对每个收益序列使用 q（=60）个数据点来计算样本平均收益 $E\hat{R}_i^{(1)} = \sum_{i=1}^{1} R_i^{(1)}/q$ 和方差协方差矩阵 $\Sigma^{(1)}$，然后以此作为输入变量来求解投资组合优化问题，我们就得到了所拟合的最优投资组合权重 $\hat{\omega}$ 及投资组合收益和方差（$\hat{R}_p$，σ_p）$^{(1)}$的第一轮值。

（4）将步骤（2）和（3）重复 m 次，使用所生成的 m 组（$\hat{R}_p$，$\hat{\sigma}_p$）值来计算平均值（“m 轮”的平均），结果记为（$\bar{R}_p$，$\bar{\sigma}_p$），从而夏普比的平均值为 $\hat{S} = \dfrac{\bar{R}_p - r}{\bar{\sigma}_p}$。将蒙特卡罗模拟值与（$R_p^*$，$\sigma_p^*$，$S^*$）的真实值进行比较，就对期望收益和协方差估计方法所产生的“偏差”进行了度量。

Jobson 和 Korkie（1980）使用20只股票进行实证研究，股票的月度数据用已知的多维分布来生成，结果表明，由拟合数据（$\bar{R}_p$，$\bar{\sigma}_p$）、已知的总体参数（R_p^*，σ_p^*）和等权重投资组合所得到的夏普比之间存在很大的差异，夏普比分别为0.08，0.34，0.27。因此，“估计误差”非常的大，对风险-收益之间的权衡关系会产生很大影响。

Frost 和 Savarino（1988）进行了类似的实证研究，结果发现，当投资组合中任意资产的权重被限制为小的正值且禁止卖空时，$\bar{R}_p - R_p^*$ 和 $\bar{\sigma}_p - \sigma_p^*$ 的“偏差”（尤其是前者）会显著地下降。另外，对风险厌恶程度是市场投资者的2倍或1/2的投资者来说（市场投资者持有“市场投资组合”（比方说），市场投资组合包含25只股票），如果均值方差优化问题中的单一证券权重都不超过3% ~5%，那么这种方法就会产生最好的结果（根据确定性等价收益来进行判断）。还要注意的是，卖空和付保证金购买都会使结果变得更糟糕。因此，只要我们对均值方差优化问题施加限制，优化问题就能得到改进（尽管改进幅度不是很大）。

使用市值权重（ω_{im}）在股票市场上进行指数追踪是相当普遍的一种做法。一个带约束的优化策略是，在最优权重（ω_i^*）与当前市值权重之差不超过当前市值权重2%（比方说）的条件下（即 $\hat{\omega}_i^* = \omega_{im} \pm 0.02\omega_{im}$）来使夏普比达到最大。Cavaglia、Melas 和 Miyashiuta（1994）发现，当投资者对当前市值权重偏离一丁点时，分散化国际投资组合（即投资者持有多个国家的证券）有助于夏普比的改进。实践中也可能出现的情形是，如果最优权重和“合理”直觉之间有很大差距，所有投资经理都不会相信这些最优权重。事实上，英国养老基金对单只股票的投资比例很少超过5%（尽管 FTSE100 指数追踪策略对 Vodaphone - Mannesmann 股票的持有比例约为15%，芬兰养老基金对诺基亚股票的持有比例超过50%）。

如果“无约束时的权重”意味着投资者对某个公司或某个产业的持有比例高得非常显著（如将90%的财富投资于小市值股票，或所投资的股票中很大一部分来自于新兴市场），那么

均值方差优化者需要对权重施加一定的约束。无约束优化问题——允许卖空——所给出的权重通常意味着，你应将大量的股票卖空，用所获得的资金来投资其他股票。实践家不会简单地相信，这样的投资策略在事后会取得成功。

也有些研究试图考察特定投资组合在统计意义上是否“接近于”均值方差优化组合（如Jobson和Korkie（1980））。然而，这些检验的功效都很差（即当均值方差方法无效时，仍倾向于不拒绝它），也没有考虑不等式约束（如不允许卖空），因此这种方法在实践中并不常用。

也值得注意的是，在最优权重的计算过程中存在一些技术问题。如果协方差矩阵很大，那么矩阵求逆就存在问题，从而最优权重对所估计的协方差矩阵的细微变化非常敏感。

学习和贝叶斯理论

投资组合理论中一个众所周知的结论是，如果在市场投资组合中增加一种定价误差为正（即Jensen $\alpha>0$）的资产，那么新投资组合的夏普比（平方）将增加（$\alpha/\sigma_\varepsilon$）^{2}，这里的$\sigma_\varepsilon^2$是市场回归模型

$$(R_i-r)=\alpha_i+\beta_i(R_m-r)+\varepsilon_i$$

中的残差。Pastor（2000）提出的一个问题，如果α存在不确定性，那么这个结论是否有意义。在单期框架中，Pastor假设投资者选择风险资产来使期末财富的期望值达到最大：

$$\max_\alpha\int U(W_{+1})\,p(R_{+1}\mid\Lambda)\,\mathrm{d}R_{+1}$$

其中，R是投资组合的收益，$p(\cdot\mid\cdot)$是概率密度函数（$p.d.f.$），Λ是给定的信息集合。然而，Pastor假设预测方程中的未知参数θ是确定的，投资者使用贝叶斯规则（见第2章）来预测密度函数$p(R_{+1}\mid\Lambda)$：

$$p(R_{+1}\mid\Lambda)=\int p(R_{+1}\mid\theta,\Lambda)\,p(\theta\mid\Lambda)\,d\theta$$

$p(\theta\mid\Lambda)$是θ的后验分布，它反过来又与先验分布$p(\theta)$和似然函数$L(\theta\mid\Lambda)$成比例：

$$p(\theta\mid\Lambda)\propto p(\theta)\,L(\theta\mid\Lambda)$$

通常的相切条件所给出的最优投资组合权重为（即最大化夏普比）

$$z_t^*=\Omega^{-1}(\mathrm{E}_tR_{+1}-r)$$

$$\omega_i^*=\frac{z_i^*}{\sum_i z_i}$$

Ω和$\mathrm{E}R_{+1}$是预测密度$p(R_{+1}\mid\Lambda)$的前二阶矩。获得ω_i^*（$i=1,2,\cdots,n$）的标准方法是只使用根据历史数据所计算的Ω和$\mathrm{E}R$的样本估计值。贝叶斯方法使用基于预测分布$p(R_{+1}\mid\Lambda)$的估计值，它考虑了样本估计值和投资者关于参数（Jensen）α和β的先验看法。$\theta=(\alpha,\beta)$的后验值是先验估计值θ_0和样本估计值$\hat{\theta}$（使用极大似然法）的加权平均。对贝叶斯方法而言，任意“新”资产的最优权重都与$\tilde{\alpha}/\tilde{\sigma}^2$、$\alpha$和$\sigma^2$（回归方差）的先验平均值成比例。

Pastor发现，只使用Ω和$\mathrm{E}R-r$的样本数据来确定ω_i^*的标准方法对美国投资者而言意味着，国外股票的最优权重约为40%。但是，由于期望收益估计的不确定性，检验结果不能拒绝零假设：最优投资组合中非美国股票的权重为零（以上内容参见Britten-Jones）。在贝叶斯理论框架下，国内股票偏好取决于国内CAPM的置信度。Pastor发现，为了解释美国居民实际的国外股票持有比例（总财富的8%），国外投资组合的先念定价误差必须在-2%~2%（每年）之间（即$\sigma_\alpha=1\%$）——因此，为了解释美国投资者的国内股票偏好，他们对国内（美国市场）投资组合效率的信心必须足够的强。

余下的情形是，在实践中，由于均值方差优化方法的优美性，它仅仅是一种投资组合配置决策方法。其他判断因素，如政治风险、不同国家或产业所处的经济周期等，与纯粹的均值方差分析一样起着重要作用。由S&P500指数所体现的当前市值权重和由无约束均值方差分析所得到的权重不同（如，人们使用样本平均值来预测平均收益和协方差矩阵）。因此，实践中倾

向于使用均值方差分析方法来考察 R_i 和 σ_{ij} 的“新预测值”是否有助于夏普比的改进。夏普比的敏感性分析通常是通过对收益和协方差这两个重要变量施加“特殊约束”而不是使用历史平均值来进行的。随着情形的变化，如果最优权重 ω_i^* 在无约束优化问题中有很大的变化，那么我们就应该对一些 ω_i^* 施加约束条件（如与目前的市值“跟踪指数”权重相比，新的最优权重不能有太大的变化，或者禁止卖空）。

总之，我们的结论是，如果下述条件成立，那么均值方差优化方法非常有用：

* 将投资组合的权重限定在某个范围之内（例如，所持有的任何单一资产在投资组合价值中的比例都不超过5%，或在“市场指数权重” ±2%的范围内取值，或禁止卖空）；

* 用收益 R_i 和协方差 $\{\sigma_{ij}\}$ 更好的预测值来替代历史平均值；

* 为了进行敏感性分析，所使用的资产数目要小（例如，使用均值方差方法来对20个（比方说）国家的指数进行资源配置）。

6.4 小结

这一章讨论了与分散化国际投资组合相关的许多理论和实践问题，要点如下：

* 单期均值方差优化问题中的数学方法可以用来计算最优投资组合权重，在代表性投资者的假设下，这种方法还能确定均衡预期收益——CAPM 模型。它也是其他更为复杂的投资组合优化技巧的基础（例如，对投资组合权重施加约束时的优化方法、第15和16章中的跨期模型）。

* 当投资组合的资产数目较小（约为25种）时，分散化国际投资组合在降低投资组合风险方面有着实际意义。分散化国际投资策略，即使随机地选择股票，也能大大降低风险。

* 当我们同时考虑期望收益和风险时，分散化国际投资通常能改进风险－收益之间的权衡关系（将有效前沿向左移动），在使用远期、期货或期权来对国外收益进行套期保值后，改进效果尤其显著。如果没有对收益进行套期保值，那么分散化国际投资的改进效果存在很大的争论，对特殊的“国内货币”投资者而言，必须采用“逐一”的方式来考察分散化国际投资的效果。

* 随机选择股票，使用等权重投资组合，追踪宽泛的市场指数（如摩根斯坦利全球指数），使用均值方差优化方法，都可以实现分散化投资。

* 接受无约束均值方差分析结果的主要问题是，最优比例（权重）对期望收益的预测值和非常高或低的均值方差预测值非常敏感。于是，在实践中，最优比例要受到某种形式的约束（例如，与当前投资比例之间的差异不能超过2%）。

* 如果考虑分散化投资组合中的不确定性，那么国内偏好问题就不会那么严重，但为什么分散化国际投资是如此的少见，这仍然是一个“谜”。

附录 I：有效前沿和资本市场线

有效前沿

我们在正文中用两基金定理推导出了有效前沿，一旦得到前沿上两个点的最优权重，通过加权平均的方法（$\omega^{(q)} = \alpha\omega^{(1)} + (1-\alpha)\omega^{(2)}$）就能画出整个有效前沿。在这个附录中，我们根据一些“原始的输入变量”——$n \times n$ 的收益协方差矩阵 Ω、$n \times 1$ 的风险资产期望收益 ER、所选择投资组合的期望收益 $ER_p = \mu_p$——来推导最优权重。优化问题为，在 $\omega' e = 1$（e 为单位向量）的约束下，在投资组合期望收益 $\omega' ER = \mu_p$ 给定的条件下，选择 ω（$n \times 1$）来使投资组

合的方差 $\sigma_p^2 \equiv \omega'\Omega\omega$ 达到最小。拉格朗日函数为：

$$\theta = \frac{1}{2}(\omega'\Omega\omega) - \lambda(\omega' ER - \mu_p) - \psi(\omega' e - 1) \tag{A1}$$

一阶条件为：

$$\Omega\omega - \lambda ER - \psi e = 0 \tag{A2}$$

$$\omega' ER = \mu_p \tag{A3a}$$

$$\omega' e = 1 \tag{A3b}$$

由（A2）就得到

$$\omega = \Omega^{-1}[\lambda ER + \psi e] \tag{A3c}$$

因此

$$ER'\omega = ER'\Omega^{-1}[\lambda ER + \psi e] = \mu_p \tag{A4}$$

$$e'\omega = e'\Omega^{-1}[\lambda ER + \psi e] = 1 \tag{A5}$$

（A4）和（A5）是关于未知数 λ 和 ψ 的两个方程，可以写为

$$\begin{bmatrix} A & B \\ B & C \end{bmatrix}\begin{bmatrix} \lambda \\ \psi \end{bmatrix} = \begin{bmatrix} \mu_p \\ 1 \end{bmatrix} \tag{A6}$$

其中，

$$A = (ER')\Omega^{-1}(ER) \tag{A7a}$$

$$B = (ER')\Omega^{-1}e \tag{A7b}$$

$$C = e'\Omega^{-1}e \tag{A7c}$$

注意，A、B 和 C 都是标量。根据（A6），拉格朗日乘子为

$$\lambda = \frac{C\mu_p - B}{AC - B^2} \tag{A8a}$$

$$\psi = \frac{A - B\mu_p}{AC - B^2} \tag{A8b}$$

将（A8a）和（A8b）代入（A3c）中，

$$\omega^* = \Omega^{-1}\left[\frac{ER(C\mu_p - B) + e(A - B\mu_p)}{(AC - B^2)}\right] \tag{A9}$$

这里的 ω^* 是风险资产的最优持有比例向量（$n \times 1$）。使用（A9），最小投资组合方差为（对给定的 μ_p 而言）

$$\mathrm{Var}(R_p) = \omega'\Omega^{-1}\omega = \frac{C\mu_p^2 - 2B\mu_p + A}{(AC - B^2)} \tag{A10}$$

投资组合的方差是投资组合平均收益的二次函数。在均值方差空间中，函数图像是一条抛物线；在均值—标准差空间中，函数图像为双曲线。

在 $\omega' e = 1$ 的约束条件下最小化 $\omega'\Omega\omega$，或直接根据（A10），在 μ_p 给定的条件最小化 $\mathrm{Var}(R_p)$，都可以得到全局最小方差投资组合：

$$\mu_p^{gmv} = B/C \tag{A11}$$

因此，全局最小方差（gmv）投资组合中的权重（使用（A9））和投资组合的方差分别为

$$\omega_{gmv}^* = \Omega^{-1}e/(e'\Omega^{-1}e) \tag{A12a}$$

$$\sigma_{gmv}^2 = 1/C \tag{A12b}$$

当然，ω_{gmv}^* 与期望收益相独立，全局最小方差投资组合确实能获得正的期望收益 $\mu_p^{gmv} > 0$，因为 C 必须是正定的二次函数形式。

也要注意的是，根据（A9），有效前沿上投资组合的权重是线性函数，因此，对有效前沿上的任意两点而言，第三个投资组合的期望收益为

$$\mu_p^{(3)} = \alpha\mu_p^{(1)} + (1 - \alpha)\mu_p^{(2)}$$

$$\omega_p^* = \alpha\omega_1^* + (1-\alpha)\,\omega_2^*$$

这就是两基金定理，有的时候也将其表述为：均值方差前沿可以由任意两个前沿收益来张成（对 CAPM 模型进行检验时，大量的文献使用这个术语）。

资本市场线

当我们引入无风险资产时，使投资组合方差最小的 Var（R_p）和 μ_p 之间存在线性关系。优化问题是，选择 ω 来最小化 $\omega'\Omega\omega$，约束条件为

$$R_p = \omega' R + \omega_f r \tag{A13}$$

这里的 ω 为 $n \times 1$ 列向量，$\omega' e + \omega_f = 1$，因此

$$R_p = \omega'(R - r \cdot e) + r \tag{A14}$$

拉格朗日函数为：

$$\theta = \frac{1}{2}\omega'\Omega\omega + \lambda\,[\mu_p - (\omega' ER - r \cdot e) - r] \tag{A15}$$

μ_p 是固定的（实现给定）。优化问题的解为：

$$\omega^* = k_p z \tag{A16a}$$

$$z = \Omega^{-1}(ER - r \cdot e) \tag{A16b}$$

$$\begin{aligned} k_p &= (\mu_p - r)\,/\,(A - 2Br + Cr^2) \\ &= (\mu_p - r)\,/\,[(ER - r \cdot e)\,\Omega^{-1}\,(ER - r \cdot e)] \end{aligned} \tag{A16c}$$

$$\omega_f^* = 1 - e'\omega^* \tag{A16d}$$

$$\mathrm{Var}(R_p) = \omega^{*\prime}\Omega^{-1}\omega^* = \frac{(\mu_p - r)^2}{(A - 2Br + Cr^2)} \tag{A17}$$

注意，k_p 是 μ_p 的线性函数（ω^* 也是 μ_p 的线性函数）。投资组合的标准差 $\sqrt{\mathrm{Var}\,(R_p)}$ 也是 μ_p 的线性函数，因此资本市场线（即在风险资产和无风险资产之间进行选择时的最优权重）在（σ_p，μ_p）空间上是一条直线。有时也称资本市场线为“存在无风险资产时的有效前沿”。

由资本市场线可以得到的另外一个漂亮结果是切点投资组合。下面来一步步地推导切点投资组合的求解过程，我们知道它位于资本市场线上。根据（A16）可知，对任意给定的 μ_p，ω^* 的解包括两个部分：k_p 和 z，k_p 依赖于 μ_p 的选择，z 与 μ_p 的选择无关。切点投资组合是资本市场线上的一个特殊点，即资本市场线与有效前沿的切点，切点投资组合与 μ_p 的选择无关。因此切点投资组合中风险资产的权重与 z 成比例，权重的和为 1。“切点”或市场投资组合中风险资产的权重 ω_q 为

$$\omega_q = \frac{z}{e'z} = \frac{\Omega^{-1}(ER - r \cdot e)}{e'\Omega^{-1}(ER - r \cdot e)} \tag{A18}$$

权重的和为 1，即 $e'\omega_q = 1$。无论投资者的风险厌恶程度如何，市场投资组合的权重 ω_q 对所有投资者而言都相同。无风险资产的权重为 $\omega_f^* = 1 - e'\omega^*$，就目前而言我们还无法确定它的值，因为我们只能确定 ω_q 的值，但不能决定 k_p 的值，因此也就无法确定 ω^* 的值（见（A16a））。这并不令人感到奇怪，因为以无风险利率借入或贷出多少资金的选择问题，确实取决于投资者的风险偏好，我们还没有引入“度量”风险厌恶程度的客观（效用）函数。只有引入一个函数之后——函数能说明个体投资者如何在风险 σ_p^2 和收益 μ_p 之间进行权衡，才能确定持有无风险资产和风险资产的最优比例。

附录 II：市场投资组合

假设允许卖空。因为$\sum_i \omega_i = 1$，所以可以将最大值函数写为：

$$\max\theta = \frac{ER_p - r}{\sigma_p} = \frac{\sum \omega_i (ER_i - r)}{\sigma_p} \tag{A1}$$

其中，

$$\sigma_p = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} \omega_i\omega_j\sigma_{ij}\right)^{1/2} \tag{A2}$$

我们先不用矩阵符号来阐述求解过程。因为ER_p和σ_p都取决于ω_i，所以方程（A1）的微分要使用“乘积法则”（即 d（uv）$=udv+vdu$）：

$$\begin{aligned}\frac{\partial\theta}{\partial\omega_i} &= \sum \omega_i\ (ER_i - r)\ \left[\frac{\partial}{\partial\omega_i}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_i\omega_j\sigma_{ij}\right)^{1/2}\right] \\ &\quad + \frac{1}{\sigma_p}\cdot\frac{\partial}{\partial\omega_i}\left[\sum_{i=1}^{n}\omega_i\ (ER_i - r)\right] \\ &= \sum \omega_i\ (ER_i - r)\ \left[\left(-\frac{1}{2}\right)\ (\sigma_p^{-3})\ \sum_{j=1}^{n}(2\omega_j\sigma_{ij})\right] \\ &\quad + \frac{1}{\sigma_p}\ (ER_i - r) \\ &= 0\end{aligned} \tag{A3}$$

两边同时乘以σ_p并注意$\sum_i \omega_i ER_i = ER_p$就得到

$$\left[\frac{ER_p - r}{\sigma_p^2}\right]\sum_{j=1}^{n}\omega_j\sigma_{ij} = (ER_i - r) \qquad (i = 1,\ 2,\ \cdots,\ n) \tag{A4}$$

在最大值处，方括号中的项为常数，我们将它记为

$$\eta = \left[\frac{ER_p - r}{\sigma_p^2}\right] \tag{A5}$$

η 通常称为 η 风险的市场价格。将方程（A5）代入（A4），我们就得到市场投资组合的一阶条件（正文中的方程（19））

$$\sum_{j=1}^{n}\sigma_{ij}z_j = ER_i - r \qquad (i = 1,\ 2,\ \cdots,\ n) \tag{19}$$

其中，$z_j = \eta\omega_j$。

求解：矩阵符号

现在，我们用矩阵符号来重复上述求解过程：

$$\theta = \frac{\omega'(ER - r\cdot e)}{(\omega'\Omega\omega)^{1/2}} = \frac{\sum \omega_i (ER - r)}{\sigma_p} \tag{A6}$$

其中，$\omega = (\omega_1,\ \omega_2,\ \cdots\omega_n)$，$ER - r\cdot e = (ER_1 - r,\ ER_2 - r,\ \cdots,\ ER_n - r)$，$\sigma_p = (\omega'\Omega\omega)^{1/2}$，$\Omega_{n\times n} = \{\sigma_{ij}\}$，$e$ 是 $n\times 1$ 单位向量。使用微分的乘积法则和链式法则就得到，

$$\begin{aligned}\frac{\mathrm{d}\theta}{\mathrm{d}\omega} &= (ER - r\cdot e)\ (\omega'\Omega\omega)^{-1/2} \\ &\quad - \frac{1}{2}\omega'\ (ER - r\cdot e)\ (\omega'\Omega\omega)^{-3/2}\ (2\Omega\omega)\ = 0\end{aligned} \tag{A7}$$

其中，$\dfrac{\mathrm{d}\ (\omega'\Omega\omega)}{\mathrm{d}\omega} = 2\Omega\omega$。两边同时乘以 $(\omega'\Omega\omega)^{1/2}$ 就得到

$$ER - r = [\omega' (ER - r \cdot e) / (\omega' \Omega \omega)] \Omega \omega \qquad (A8)$$

方括号中的项是一个数，等于投资组合的超额收益（$ER_p - r$）与组合的方差 $\sigma_p^2 = (\omega' \Omega \omega)$ 的商，对任意一组 ω_i 而言它都是常数。我们将这个常数记为 η（参见方程（A5）），（A8）变为：

$$ER - r \cdot e = \Omega (\eta \omega) = \Omega z \qquad (A9)$$

其中，$z = \eta \omega$。因此，解为 $n \times 1$ 向量 z：

$$z = \Omega^{-1} (ER - r \cdot e) \qquad (A10)$$

最优市场投资组合的权重为：

$$\omega_q = \frac{z}{e'z} = \frac{\Omega^{-1} (ER - r \cdot e)}{e' \Omega^{-1} (ER - r \cdot e)}$$

因为 $\sum_i \omega_i = 1$，所以每一种资产的权重为 $\omega_{q,i} = z_i / \sum_i z_i$。这就完成了最优“切点”组合或“市场投资组合”权重 z（严格地说是 ω）的推导。

第7章

绩效度量、CAPM与APT

目的

- 定义一些概念来对投资组合管理人员的投资绩效进行排序，考察他们是否能“击败市场”。
- 放松标准CAPM模型的一些严格假设。
- 说明单因子模型（SIM）的有用性。
- 考察套利定价理论（APT），它是基于“无套利”的多因子模型。

7.1　绩效度量

对不同交易者或同一交易者在不同时期的实际投资绩效进行评估是非常有用的。绩效度量指标包括相对于投资组合风险而言的收益，根据这个指标可以对投资组合进行相应排序。能够获得比其他交易者收益都高的交易者并不必然是“最好”的交易者，因为他的投资组合可能处于一个更高的风险水平。“主动型”共同基金允许基金经理对宽泛的证券组合进行投资，基金经理试图选取“赢家”（winners）。绩效指标很有用的一个领域是对“主动型”共同基金进行排序，考察它们的绩效是否超过“被动型”投资策略（如对指数进行追踪）。我们所考察的绩效度量指标有夏普、Treynor和Jensen等指标。

例如，假设基金经理的绩效评估是每个月进行一次。在5年的时期中，假设两位基金经理A和B的“主动型”投资组合策略的（超额）收益相同，每年都为10%。如果基金经理A的月收益（用年利率表示）分别为10.1%，9.8%，10.2%，9.9%，…，而基金经理B的月收益分别为20%，2%，25%，0%，-39%，…，那么你最满意的基金经理是哪一位呢？如果你的投资时期只有一个月，你喜欢收益厌恶风险（即标准差），那么你肯定会选择基金经理A。这个简单例子说明，绝大多数投资者不仅关心特殊投资策略的平均收益，还关心投资策略的风险。我们还没有明确的风险度量指标。在上面的例子中，显然，基金经理B收益的标准差比基金经理A高，基金经理A的夏普比比基金经理B大——夏普比度量的是“每单位风险的平均收益”。

什么原因导致了基金经理A和基金经理B收益的不同变化呢？也许是因为基金经理B的投资组合有着较高的β值，而基金经理A的投资组合的β值却较低。因为两位基金经理都获得了相同的平均收益，所以我们推测，基金经理B每单位“β风险”的平均收益比基金经理A低。在这里，我们用投资组合的β值作为两位基金经理投资组合（相对）风险的度量——这是Treynor绩效指标的基础。最后要注意的是，根据CAPM模型，每位基金经理（$i=A$或B）

的“必要收益”为 $RR_i=\beta_i\ (\bar{R}_{m,t}-\bar{r})$。因此，对经 CAPM 风险调整后的必要收益而言，基金经理 A 比基金经理 B 低，因为基金经理 A 的股票组合有着更低的 β 值。对每一位基金经理而言，平均超额（异常）收益 α_i 等于实际的历史平均超额收益减去 RR_i。这就是 Jensen “α”，它也是一个广泛使用的、经风险调整后的绩效度量指标，这种度量方法可以扩展到多因子模型。

夏普比

夏普所建议的度量指标是“报酬－方差比率”，对投资组合而言，夏普比定义为

$$SR_i=\frac{ER_i-r}{\sigma_i} \tag{1a}$$

其中，ER_i 是投资组合的期望收益，σ_i 是投资组合收益的方差，r 是无风险收益率。夏普比可以用来度量变换线的斜率，正如我们已经看到的那样，在均值－方差组合理论中，我们可以用夏普比来选择风险资产的最优投资权重（允许借贷）——“切点投资组合”。按照这种方式，我们可以将（1a）描述为事前的夏普比。基于投资组合期望收益和方差预测值（如使用 EWMA 或 GARCH 模型）的事前夏普比可以用来对你将要（比方说，明年）持有的各种投资组合进行排序。实践中主要根据事后夏普比来对投资组合进行排序，事后夏普比是基于历史的“报酬－方差比率”：

$$SR_i=\frac{\bar{R}_i-\bar{r}}{\sigma_i} \tag{1b}$$

使用投资组合的（比方说）月度收益数据（和无风险收益率一起使用）就能计算出任何风险组合的夏普比。历史夏普比最大的投资组合就是“最好”的投资组合。夏普比被广泛用于各种“主动型”策略（如“股票选择”策略或“择时交易”策略）和“被动型”策略（如对 S&P500 指数进行追踪的策略）的比较，也广泛用于各种基金的绩效比较。人们通常用夏普比来对两个（或更多）良好的分散化投资组合进行比较，假设条件是，你只对夏普比最大的“最好投资组合”进行投资。

正如下面将看到的那样，尽管夏普比被广泛使用，其背后的隐含假设有着非常强的限制性。当我们对各种投资组合的历史绩效进行比较时，如果下述条件成立，那么“初始”夏普比就是“正确”的统计指标：

（ⅰ）初始时刻的投资组合只包括无风险资产；

（ⅱ）然后，考虑两种（或更多）相互排斥的零投资策略（zero－investment strategy），将借来的资金投资于基金 X 或基金 Y（或其他基金）。

零投资策略包括无风险资产的空方头寸（即借款）和风险资产的多方头寸（Sharpe（1966，1975））。当我们使用夏普比来对投资组合进行排序时，一个关键的假设是，所选择的基金（$F=X$ 或 Y）与现有投资组合的相关系数为 0——当现有投资组合为“现金”（即无风险资产）时，这一点就可以得到保证；但是，如果初始投资组合包含风险资产，那么这一点就无法得到保证——正如我们在下面将看到的那样。

情形 A：现有投资组合为“现金”

为了说明在条件（ⅰ）和（ⅱ）成立时为什么“初始”夏普比是有效的，考虑如下情形。假设你在目前的投资组合中投入 A 美元的资金——它是无风险资产，能获得的 r 收益。现在，你借入 V 美元来对风险基金 F 进行投资，因此，在 $t+1$ 时刻，你的“新”投资组合（即原有的现金加上 F 中的零投资策略）的价值为：

$$\begin{aligned}W_{t+1}&=A\ (1+r)\ +V\ [\ (1+R_F)\ -\ (1+r)]\\&=A\ (1+r)\ +V\ (R_F-r)\end{aligned} \tag{2}$$

“新”投资组合的百分比收益（以初始的 A 美元作为计数单位）为

$$R_p = r + (V/A)(R_F - r) \tag{3}$$

这里的 V/A 是杠杆比率。于是

$$\sigma_p = (V/A)\sigma_F \tag{4}$$

根据定义，基金 F 的夏普比为

$$SR_F = (R_F - r)/\sigma_F$$

为了简单起见，我们对期望符号进行了简化。于是

$$R_p = r + (V/A)\sigma_F SR_F \tag{5}$$

于是，当 σ_F 已知时，投资者选择 V 来使投资组合的风险 σ_p^* 达到合意水平，从而根据（4）有 $V^* = A\sigma_p^*/\sigma_F$。但是，就目前而言，当投资者的风险承受水平 σ_p^* 给定时，投资者根据最大夏普比来选择基金 F，然后使新投资组合（参见（5））的期望收益达到最大。这就是根据最大夏普比来对基金 F 进行零投资策略选择的基础，但这个“规则”依赖于“现有投资组合是无风险资产”这一事实。

出于完整性考虑，值得注意的是，与夏普比非常接近的概念是信息比率（information ratio, IR）：

$$IR_i = R_i/\sigma_i \tag{6}$$

相对于夏普比而言，IR 只是“省略了”分子中的无风险利率。然而，对两只基金 A 和 B 而言，即使 $SR_A < SR_B$，$IR_A > IR_B$ 也是可能的，因此两种标准所给出的排序不同。夏普比是所使用的“正确”统计量，因为它表示零投资策略的收益，具有易于计算的优点。

情形 B：初始投资组合包含风险资产

现在，我们使情况稍微复杂一些，假设现有投资组合不是“现金”而是包含风险资产的投资组合 E，由此来对“原始”夏普比进行变动。事先存在的投资组合 E 包含风险资产，你试图在现有投资组合 E 中增加零投资组合，两个（或更多）零投资组合之间相互排斥，风险资产与这些零投资组合相关。另外一种复杂情况是，在实践中，我们经常希望将所选择投资组合的收益与某个基准投资组合（如 S&P500）的收益（而不是无风险利率）进行比较。对即将加入现有投资组合的各种选择而言，我们仍然能用夏普比来对他们进行排序吗？如果我们非常的小心，那么答案是肯定的。我们希望将基金加入到现有投资组合 E 中去，基金 i（相对于基准风险投资组合而言）的“新”夏普比为

$$SR_i = \frac{R_i - R_b}{\sigma(R_i - R_b)} = \frac{R_d}{\sigma_d(R_d)} \tag{7}$$

其中，$R_d = R_i - R_b$ 是所选择基金和基准投资组合之间的收益差。当 $R_b = r$ 时，SR_i 就是初始夏普比。假设你可以选择将基金 X 或基金 Z 加入到现有风险组合 E 中去，且 $R_X > R_Z$。它们的夏普比分别为：

$$SR_X = \frac{R_X - R_b}{\sigma(R_X - R_b)} > SR_z = \frac{R_z - R_b}{\sigma(R_z - R_b)} \tag{8}$$

你的新投资组合的夏普比要么为 SR_{E+X}，要么为 SR_{E+Z}。如果 $SR_X > SR_Z$，那么是否总有 $SR_{E+X} > SR_{E+Z}$？答案是否定的。为了直观地看清这一点，假设现有投资组合 E 的收益与基准投资组合的收益 R_b 无关。现在，为了简化起见，假设投资组合 E 和基金 X 的收益正相关，而投资组合 E 和基金 Z 的收益负相关。将 X 加入到 E 中，新投资组合可以获得较大的超额收益，$R_X > R_Z$，但新投资组合的风险也有所增加，因为 $\rho_{EX} > 0$。将 Z 加入到投资组合 E 中就不可能获得这样的超额收益，但可以降低新投资组合的风险。因此，我们不能肯定地说，$SR_{E+X} > SR_{E+Z}$。我们必须做的就是直接计算两种可能的新投资组合的夏普比。

为了看清计算是如何进行的，假设初始时刻你在现有投资组合上的投资为 A 美元，你用卖空基准投资组合的资金来购买基金 F（X 或 Z，但不是两者的组合）。我们再一次对零投资组合的夏普比进行定义，于是

$$W_{t+1} = A\ (1 + R_E)\ + V\ (R_F - R_b) \tag{9a}$$

$$R_{E+F} \equiv\ (W_{t+1}/A)\ - 1$$

$$= R_E +\ (V/A)\ (R_F - R_b) \tag{9b}$$

$$\sigma^2_{E+F} = \sigma^2\ (R_E)\ +\ (V/A)^2\sigma^2\ (R_F - R_b)$$

$$= +2\ (V/A)\ \sigma\ (R_E)\ \sigma\ (R_F - R_b)\ \rho\ (R_E,\ R_F - R_b) \tag{9c}$$

因此，“新”投资组合的标准差（E 加上 F 的多方头寸和基准组合的空头头寸）取决于 R_E 和（$R_F - R_b$）之间的相关系数，在计算夏普比 SR_{E+F}时（$F = X$ 或 Z）要用到这个相关系数。SR_{E+F} 可以表示为：

$$SR_{E+F} = \frac{R_{E+F} - R_b}{\sigma\ (R_{E+F} - R_b)} \tag{10}$$

我们仍然可以使用以无风险利率为基准投资组合情形中所勾画的程序。投资者选择风险承受水平 σ^*_{E+F}，根据（9c）来计算 V^*，然后根据（9b）来确定 R_{E+F}的大小。最后，计算夏普比 SR_{E+F}。当 $F = X$ 或 $F = Z$ 且 $\sigma^*_{E+X} = \sigma^*_{E+Z}$时，上述程序同样适用，从而两种选择的风险承受水平相同。然后，基于各自的夏普比——使用（10）来进行计算并假设 $SR_{E+F} > SR_E > 0$，就可以选择将 X 或者 Z 加入到现有投资组合 E 中去。

情形 C：新旧投资组合

上述讨论说明，当投资组合包含风险资产时，我们不得不计算“新”投资组合——E 加上所考虑的新增资产 F（等于 X 或 Z）——的夏普比。为了看清这一结论背后的直觉，我们对符号进行简化，用 $E + F$ 表示“新”投资组合。于是，夏普比意味着，如果

$$SR^{new} > SR^E \quad \text{或} \quad \frac{R^{new} - R_b}{\sigma^{new}_d} > \frac{R_E - R_b}{\sigma^E_d} \tag{11a}$$

那么我们就对新投资组合进行投资（Dowd（2000）），其中

$$R^{new} = zR_z +\ (1 - z)\ R_E = z\ (R_z - R_E)\ + R_E \tag{11b}$$

R^{new}是可供选择的风险投资组合 Z 的收益与现有投资组合 E 收益的一个线性组合。注意，$\sigma^E_d \equiv \sigma\ (R_E - R_b)$ 已经包含了 R_E 和 R_b 之间的任何相关性，而 $\sigma^{new}_d \equiv \sigma\ (R^{new} - R_b)$ 则包含了 R^{new}和 R_b 之间的任何相关性。

重新排列（11a）并使用（11b），很容易证明，夏普比准则意味着

$$R_z - R_b >\ (R_E - R_b)\ + \frac{(R_E - R_b)}{z}\left(\frac{\sigma^{new}_d}{\sigma^E_d} - 1\right) \tag{12}$$

根据（12）可以得出一些直观结论：

（a）如果 $\sigma^{new}_d = \sigma^E_d$，那么只要 $R_Z > R_E$，你就会把投资组合 Z 增加到现有投资组合中去；

（b）如果 $\sigma^{new}_d > \sigma^E_d$，那么为了将投资组合增加到现有投资组合中去（也假设 $R_E - R_b > 0$），我们需要一个充分大的 R_z（$> R_E$）去弥补投资组合更高的整体风险。

夏普比与一些效用函数存在联系，这些效用函数是期末财富的函数。对小的对称风险而言，所有效用函数的行为都与二次函数（台劳级数的二阶展开式）相似。SR 与二次效用函数紧密相关，基于（期望）二次效用最大化的投资组合选择，与基于（存在无风险资产的情形）夏普比最大化的投资组合选择之间存在一一对应关系。然而，如果财富变化导致投资越过了极乐点，或收益具有肥尾或偏斜特征，那么夏普比就不是一个很好的“回报－风险”度量指标。为了避免这些陷阱，Hodges（1998）在指数效用函数（即常绝对风险规避型效用函数）的假设条件下将夏普比进行了扩展，使其适用于那些追求期望效用——期末财富的函数——最大化的投资者。

Treynor 比率

Treynor 比率（1965）的定义是：

$$TR_i = \frac{ER_i - r}{\beta_i} \tag{13}$$

因此，它是（事前）每单位风险超额收益的一种度量，但这里的风险是用投资组合的 β 值来进行度量的。如果 CAPM 模型成立，那么

$$\left(\frac{ER_i - r}{\beta_i}\right) = ER_m - r \tag{14}$$

所有证券组合的 TR_i 值都相同。与夏普比一样，可以用 Treynor 比率来对各种投资组合（投资策略）的历史绩效进行比较，“最好”的投资组合就是 Treynor 比率最大的投资组合。

当方程（14）的右边部分给定时，我们能计算出市场投资组合超额收益的平均值 $\bar{R}_m - \bar{r}$，将 $(R_i - r)_t$ 对 $(R_m - r)_t$ 进行时间序列回归，我们也能估计出任意投资组合的 β_i 值。当投资组合的平均超额收益 $\bar{R}_i - \bar{r}$ 给定时，我们能计算出（13）中每一部分的值。如果“主动型”基金管理人员的某个特殊投资组合（记为 i）的 TR_i 值大于市场投资组合的平均收益 $\bar{R}_m - \bar{r}$，那么“主动型”基金管理人员的绩效比所选择的（被动型）市场投资组合（如 S&P500）的绩效好。如果 $\beta_i < 0$，那么在解释 Treynor 比率时就存在一些困难，尽管如此，Treynor 比率也可以用于对各种风险投资组合进行排序。$\beta_i < 0$ 的情形在实践中并不常见（我们不会深入考虑这种情形）。

Treynor 比率也可以表示成回归方程的形式：

$$R_{i,t} - r_t = \alpha_i + \beta_i \ (R_m - r)_t + \varepsilon_{i,t}. \tag{15}$$

有时也将 Treynor 比率定义为 $TR_i^* = \alpha_i/\beta_i$。于是，如果

$$\alpha_A/\beta_A > \alpha_B/\beta_B \tag{16}$$

那么我们将选择基金 A 而不选择基金 B。将（16）代入（15），最大化 $TR_i^* = \alpha_i/\beta_i$ 的 Treynor 准则所给出的投资组合排序，与基于 $TR_i = (\bar{R}_i - \bar{r})/\beta_i$ 的投资组合排序相同。Treynor 比率所给出的排序与夏普比所给出的排序始终相同吗？如果 $TR_A > TR_B$，那么

$$\frac{\bar{R}_A - \bar{r}}{\beta_A} > \frac{\bar{R}_B - \bar{r}}{\beta_B}$$

使用 $\beta_A = \rho_{Am}\sigma_A/\sigma_m$，于是

$$SR_A/\rho_{Am} > SR_B/\rho_{Bm} \tag{17}$$

其中，$SR_i = (\bar{R}_A - \bar{r})/\sigma_A$。因此，为了使 Treynor 比率所给出的排序与夏普比所给出的排序相同，必须要求基准投资组合为无风险资产，且两种投资组合的收益与市场收益的相关性相同。显然，两种准则所给出的投资组合排序不一定相同。

Jensen α

Jensen 指数（“α”）假设投资者持有一个“主动管理型”投资组合 i，投资组合的绩效用如下回归方程中的截距 α_i 来进行度量：

$$R_{i,t+1} - r_t = \alpha_i + \beta_i \ (R_{m,t+1} - r_t) + \varepsilon_{i,t+1} \tag{18}$$

“选股者”在回归过程中可能会使用 60 个或更多月份的（超额）市场收益数据和投资组合的超额收益的数据。由方程（18）可以直接看出，如果 $\alpha_i = 0$，那么我们就得到了标准的 CAPM 模型。因此，如果 $\alpha_i > 0$，那么投资组合就会获得超过 CAPM/SML 所给出的收益。$\alpha_i > 0$ 意味着投资组合位于 SML（证券市场线）之上。当 $\alpha_i < 0$ 时，相对于投资组合的 β 风险而言，主动型投资组合管理人员的绩效要更低一些。注意，Jensen α 准则所给出的投资组合排序可能与基于夏普比的排序不同（以无风险资产作为基准）。容易证明，如果 $\alpha_A > \alpha_B$，那么只有当 $\rho_{Am} = \rho_{Bm}$ 且 $\sigma_A = \sigma_B$ 是，才会有 $SR_A > SR_B$。然而，通常（显然并不总是）出现的情形是，夏普、Jensen 和 Treynor 三种度量指标所给出排序相同，对每个投资组合而言，我们可以计算出三个绩效度量指标。

人们通常用比（18）更为复杂的一个回归方程来估计“Jensen α”，回归方程右边的变量

反映了期望收益上的可能影响（如，β_i 可能是时变的，取决于一组经济变量——如红利收益、收益差异等）。然而，$\alpha_i>0$ 时的“表现”仍然要更好一些。

我们对（18）进行重新排列

$$TR_i \equiv \frac{E_t R_{i,t+1} - r_t}{\beta_i} = \frac{\alpha_i}{\beta_i} + (E_t R_{m,t+1} - r_t) \tag{19}$$

（19）的左边就是 Treynor 比率 TR_i。如果 $\beta>0$（这对绝大多数组合而言都成立），那么我们很容易看出，当 $TR_i > E_t R^m_{t+1} - r$ 时，$\alpha_i>0$。

Roll 批评与绩效度量

Roll 批评——它关心的是使用样本数据来对 CAPM 模型进行估计时的问题——意味着，对任何数据集合而言，如下关系始终成立：

$$\bar{R}_i = \bar{r} + (\bar{R}_m - \bar{r}) \cdot \hat{\beta}_i \tag{20}$$

$\bar{R}_i$ 是投资组合 i 收益的样本均值，$\bar{R}_m$ 是市场投资组合收益的样本均值。如果市场投资组合的收益度量是正确的，那么在任何样本数据集合中，投资组合的期望收益和 β 值之间都是一种精确的线性关系。因此，如果 CAPM 模型是投资行为的正确描述，那么 Treynor 指数始终等于市场投资组合的样本超额收益，Jensen 指数始终等于 0。于是，如果使用的是 Treynor 指数或 Jensen 指数而不是 Roll 所建议的度量指标，那么这意味着我们对市场投资组合度量不正确。

面对 Roll 的批评，人们只能在非常特殊的情形中使用这些绩效指标。如果我们发现，三种绩效指标对市场投资组合（如 S&P500，NYSE 指数）的排序保持不变，那么 Treynor、夏普和 Jensen 指标就为各种投资组合的相对绩效提供了非常有用的描述统计。

7.2 CAPM 模型的扩展

CAPM 模型表明，对所有股票（和所有投资组合）而言，经“风险 β_i”调整后的期望超额收益都相同：

$$(ER_i - r)/\beta_i = (ER_j - r)/\beta_j = \cdots \tag{21}$$

方程（21）在标准 CAPM 模型的严格假设下成立，这些假设包括：

- 所有投资者有着相同的预期；
- 投资者选择投资组合权重来使夏普比达到最大；
- 投资者能以无风险利率无限制地进行借或贷；
- 市场在任何时刻都是均衡的。

对某些假设进行放松是可能的，例如使用不同的借贷利率，考虑税收、不可交易资产、（价格）通货膨胀。然而，现有文献很少讨论这些假设的“变体”。我们在此简略地讨论两个值得讨论的问题：“零 β CAPM 模型”；在均值－方差模型中引入异质型预期。

零 β CAPM 模型：无风险资产不存在的情形

尽管投资者能以无风险利率贷出任何资金（如购买政府债券），投资者通常不能无限制地借入资金。另外，如果未来通货膨胀存在不确定性，那么从实际结果（而不是名义结果）来看，无风险借入是不可能的（如果政府发行指数连动债券，那么无风险贷出仍然可能）。

这一节中在不存在无风险借或贷的假设下（仍然允许卖空）来重新阐述 CAPM 模型，由此可以得到 Black（1972）的零 β CAPM 模型。因为无风险资产不存在，所以模型决定的是任何资产 i（或者资产组合）的实际收益，

$$ER_i = ER_z + (ER_p - ER_z) \cdot \beta_{ip} \tag{22a}$$

这里的 ER_z 是所谓的零 β 组合的实际期望收益，零 β 组合与任意有效的最小方差组合 p 相伴随

（参见图1）。在与投资组合 p 无关的所有投资组合中，零 β 组合有着最小的方差。（注意，投资组合 p 不是标准CAPM模型中的“市场投资组合”。）资产 i 的 β 值为

$$\beta_{i,p}=\frac{\mathrm{Cov}\ (R_i,\ R_p)}{\mathrm{Var}\ (R_p)} \tag{22b}$$

根据构造，零 β 组合与相应的有效均值-方差组合 p 之间的协方差为零。注意，ZZ' 上所有组合都是零 β 组合，但 Z 是具有最小方差的投资组合（在这个特殊的投资组合集合内）。也可以证明，Z 始终是一个无效投资组合（即它位于有效前沿的下半支上）。

因为在有效前沿上我们可以任意选择投资组合 p，所以构造与“p”相似的组合与相应零 β 的无穷多个线性组合是完全可能的。因此，我们就损失了标准CAPM模型中所发现的一个重要性质，无论投资者的偏好如何，所有投资者都选择相同的风险资产集合。但这是一个更为现实的结果，因为个体投资者所持有的风险资产组合不同。资产 i 的均衡收益可以用（22a）来很好地表述，也可以用 P^* 和 Z^* 的线性组合来表示

$$ER_i=ER_{z^*}+\ (ER_{p^*}-ER_{z^*})\ \beta_{ip^*}^* \tag{22c}$$

当然，（22a）和（22c）所得出的资产的期望收益必须相等。这个结论与标准CAPM模型形成对比，在标准CAPM模型中，无风险资产与市场投资组合的线性组合是唯一的投资机会集合。另外，在零CAPM模型中，射线 XX' 并不表示可行投资机会集合。

在任意两个有效投资组合 p_i（$i=1,\ 2$）及与它们相对应的正交风险组合 Z_i 给定的条件下，对这两个有效投资组合进行线性组合，所有投资者都可以（不用借或贷）实现合意的最优投资组合。因此，两基金定理对零 βCAPM模型也成立。

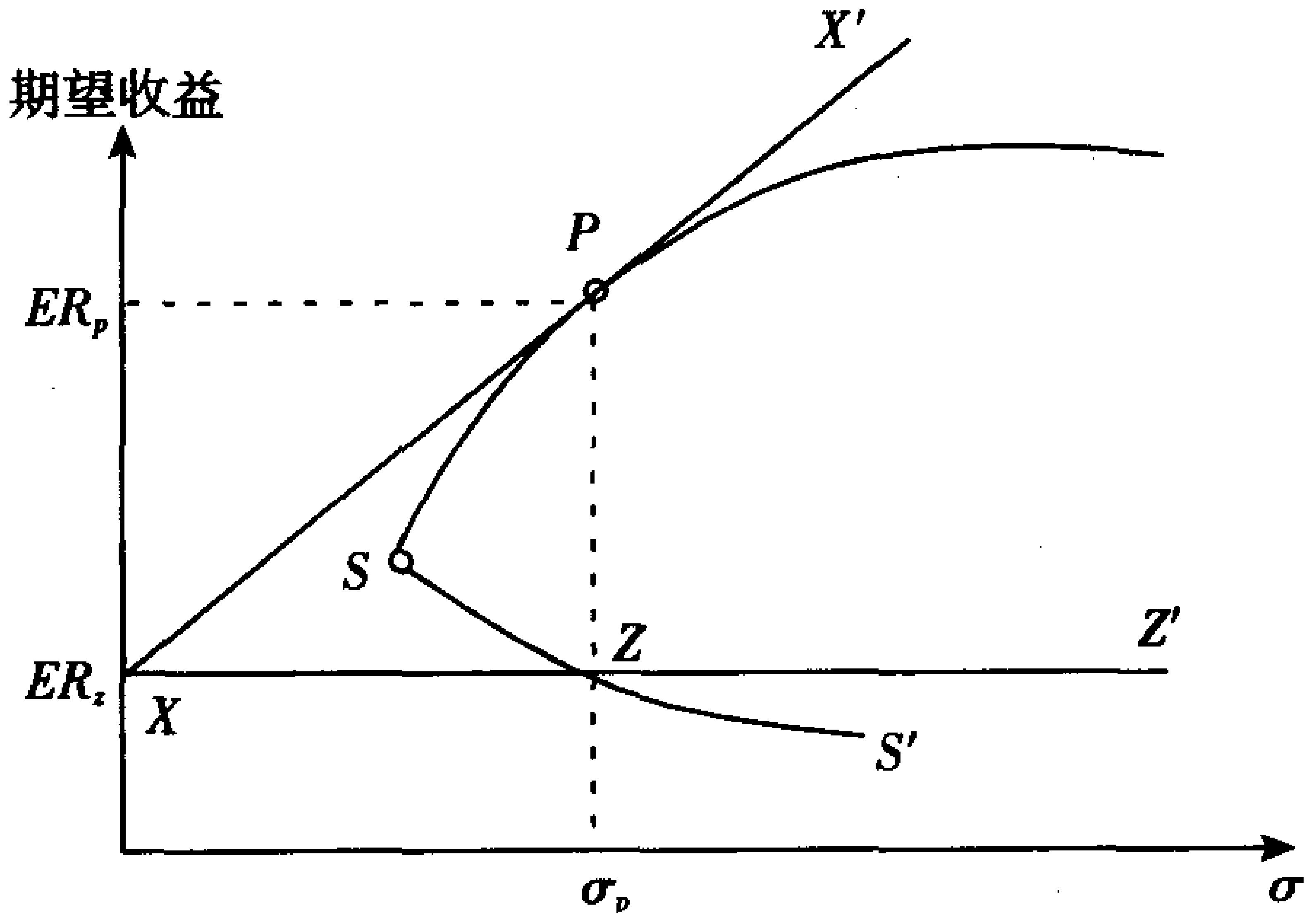


图1　零 βCAPM模型

与标准CAPM模型一样，零 βCAPM模型也是均衡收益模型，主要特征如下：

* 将均值-方差有效组合 M 与相应零 β 组合进行线性组合，个体投资者不用（以无风险利率）借或贷就可以实现最优投资组合（投资者的偏好给定）。

* 零 β 组合是由风险资产所组成的一个投资组合，期望收益为 $ER_z=\sum_i\omega_iER_i$，风险资产的权重为 ω_i。零 β 组合具有如下性质：

（ⅰ）零 β 组合与有效投资组合 p 无关；

（ⅱ）零 β 组合具有最小方差（对正交投资组合 p 而言，见图1）。

* 投资组合 p 和 z 的组合不唯一。然而，任意资产 i（或资产组合）的均衡收益是 ER_z 和 ER_p 的线性函数，如方程（22a）所示。

异质型预期

投资者对期望收益、方差和协方差的主观预期可能不同。在理性预期假设下，所有投资者在任何时点都知道随机收益的真实概率分布（模型），这种情形与异质型预期不同。在异质型预期假设下，每位投资者对有效前沿都有自己的主观预期。每位投资者所面临的问题是标准的均值方差问题，在最优资产比例的某个约束下——不同投资者的最优资产比例不同——来最大化夏普比。

总之，为了使每种资产的市场出清，对所有投资者（$k=1, 2, \cdots, p$）进行加总就得到了风险资产期望收益的表达式，它是投资者主观偏好（风险与收益）和资产协方差复杂的加权平均。边际替代率取决于个体的财富水平。因此，均衡收益和资产价格都取决于财富水平，但财富本身是由价格来决定的，所以在异质型预期情形中不存在“闭型”解或显示解。

如果对效用函数施加限制，使期望收益和风险（方差）之间的边际替代率不再是财富的函数，那么我们就能得到异质型预期情形中的解。Linter（1971）假设效用函数是财富的负指数函数，这意味着绝对风险规避系数为常数。即使在这种情形中，期望收益（尽管与财富独立）仍然取决于个体对协方差的主观预期与个体风险规避系数之间复杂的加权平均值。

7.3 单因子模型

单因子模型（SIM）不是体现行为假设（例如，与补偿“市场风险”的必要收益有关的假设）的真正“模型”，它只是一个统计假设：任意证券的收益 R_{it} 都可以充分地表示为某个单一（经济）变量的线性函数（如通货膨胀率或利率），

$$R_{it}=\theta_i+\delta_i I_t+\varepsilon_{it} \tag{23}$$

其中，ε_{it} 是随机误差项，方程（23）对任意时点上的任意证券（或投资组合）（$i=1, 2, \cdots, n$）都成立。因此，I_t 可以是与 R_{it} 相关的任意变量，单因子模型不是用来解释这些相关性的某个特殊理论模型。

显然，即使所选择的变量 I_t 是市场投资组合超额收益 R_m-r，SIM 和 CAPM 的方程形式也不同，因为 CAPM 使用的是资产的超额收益，表达式中截距项也为 0。然而，有些文献将 CAPM 和 SIM 同等对待（实际上它们是两个不相同的模型），也将 CAPM 称为“单因子模型”。当 I_t 为市场超额收益（R_m-r）且被解释变量为股票的超额收益（R_i-r）时，后一种说法是可以接受的。我们故意将 I_t 的系数记为 δ_i 而不是 β_i，目的是为了强调 SIM 和 CAPM 之间的差异。

使用60个月（比方说）的时间序列数据，用 OLS（或像 GMM 等更为复杂的技巧）就可以对方程（10）进行估计，由此可以得到参数（θ, δ）、误差项的方差 $\sigma^2_{\varepsilon,i}$ 和其他“诊断性输出结果”——如回归方程的 R^2——的估计值。严格地说，SIM 也假设，证券的收益中“预料不到的部分”——用 $\varepsilon_{i,t}$ 来表示——与证券 j 相应部分之间是相互独立的，

$$\text{Cov}(\varepsilon_{it}, \varepsilon_{jt})=0 \tag{24a}$$

（对 $i\neq j$ 和所有 t 都成立。）SIM 也假设，对所有 i 和 t，I_t 与 ε_{it} 相互独立，

$$\text{Cov}(I_t, \varepsilon_{it})=0 \tag{24b}$$

在上述假设下很容易证明：

$$\bar{R}_{it}=\theta_i+\delta_i\bar{I}_t \tag{25a}$$

$$\sigma^2_i=\delta^2_i\sigma^2_I+\sigma^2_{\varepsilon,i} \tag{25b}$$

$$\sigma_{ij}=\delta_i\delta_j\sigma^2_I \tag{25c}$$

方程（25a）表明，股票的平均收益取决于因子（指标）的平均值 $\bar{I}_t$ 和股票本身的 δ_i 和 θ_i 值。受因子影响更大的股票（即 δ_i 更高且 $\delta_i>0$）有着更高的平均收益（对 $\bar{I}_t>0$ 的情形而言）。我

们从（25*b*）可以看出，对任意单个证券而言，SIM 能将收益的总方差 σ_i^2 进行分解，因为：

（a）证券的系统风险是由因子 I_t（对所有证券而言，因子是共同的）所导致的；

（b）特殊风险 $\sigma_{\varepsilon,i}^2$ 是由那些只影响单个证券的随机事件所导致的。

于是，

证券 i 收益的总方差 = “δ” × 因子 I_t 的方差 + 特殊风险的方差

一般而言，对月度数据来说，方程（10）的 R^2 相当的低，这说明单只股票收益的大部分变动都不能用变量 I_t 来很好地进行解释。因此，当我们将市场收益作为因子 I_t 来使用时，σ_i、（$\delta_i\sigma_I$）和 $\sigma_{\varepsilon,i}$ 的“球场”（ball - park）值（每年）分别为 40%、0.9 × 15% = 13.4% 和 26.6%。这仅仅说明，个体股票月度收益的总变化（这里为 40%）主要应归因于特殊风险（这里为 26.6%）而不是市场风险（13.4%）。当然，投资组合理论指出，如果我们将证券作为（良好）分散化投资组合的一个部分来进行持有，那么几乎不用支付成本就可以将这些特殊风险分散掉。这也是如下现象的原因：CAPM/SML 预言，任何股票的平均收益不是取决于“自身的方差”，而是取决于它与投资组合中其他股票之间的协方差（相关系数）。

方程（25c）表明，在单因子模型中，任意两种证券之间的协方差仅仅取决于 δ_i、θ_i 和 σ_I^2，在计算股票组合的在险价值（Var）时，这一结论非常有用（参见第 28 章）。

实际上，对一个国家的股票收益而言，独立性假设 Cov（ε_{it}，ε_{jt}）很少成立。原因是，影响 *A* 公司收益的“冲击”或“信息”不可能不影响 *B* 公司的收益。在对不同国家之间的收益进行比较时，SIM 有着更强的实用性，因为宏观经济冲击（如预期不到的利率变化）不可能在所有国家中同时发生。然而，E（$\varepsilon_i\varepsilon_j$）= 0 假设的违背并不会导致 δ_i 和 θ_i 的估计偏差，因此，在实践中人们广泛使用 SIM（参见 Cuthbertson 和 Nitzsche（2001a））。此外，将 SIM 扩展为多因子模型，使其包含更多的影响所有股票收益的变量（影响程度更大或更小），如利率、汇率、诸如“高面值市值比减低面值市值比”（HML）股票收益（参见后面的章节）等模拟因子（factor mimiking）等宏观经济变量，这种扩展能“改进”模型的绩效。与 SML 一样，多因子模型可以用来挑选那些价值被低估或高估的股票。套利定价理论 APT（参见下一节）就是一个多因子模型，它体现了分散化投资组合对股票平均收益影响的理论思想。

7.4　套利定价理论

在确定单只股票和股票组合的期望收益方面，套利定价理论（APT）是 CAPM 的一种替代。从宽泛意义上讲，APT 意味着证券收益可以分解为预期收益和无法预期的（“奇怪的”或“消息所导致的”）收益两个部分。对单只股票而言，“消息”可以进一步分解为对所有股票都有影响的“一般性消息”和只对单只股票有影响的特殊“消息”两个部分。例如，对所有股票都有影响的消息可能是无法预料的中央银行提高利率的公告；影响某个特殊产业中股票的消息可能是新雷达系统的发明——这项发明只对航空业产生影响而不会对化工或服务等其行业产生影响。APT 认为，“一般性消息”会影响所有股票的收益，但影响的程度不同。例如，利率意想不到地上涨 1%，这对部分公司的影响较大，而对部分公司的影响却较小。从某种程度上讲，APT 是比 CAPM 更为一般的模型，因为 APT 对影响某个特殊证券收益的多种因素进行了考虑，而 CAPM 只对影响期望收益的唯一因素——证券与市场投资组合收益之间的协方差——进行了考虑。APT 可以表示为：

$$R_{it} = R_{it}^e + u_{it}$$

其中，R_{it} 是第 i 只股票的实际收益，R_{it}^e 是第 i 只股票的预期收益，u_{it} 是预料不到的收益。我们可以将预料不到的收益 u_{it} 进一步细分为系统风险或市场风险——对所有股票都或多或少地产生影响——的收益 m_t 和非系统性风险（异质型风险或特殊风险）——只对一家公司或几家公司产生影响——的收益 ε_{it} 两个部分：

$$u_{it} = m_t + \varepsilon_{it} \tag{26}$$

与 CAPM 模型中的情形一样，系统风险是无法分散的，因为这些消息或新信息对所有公司都会产生影响。但是，特殊风险是可以分散的，从而无法对它进行“定价”。

为了使 APT 更有操作性，我们需要一些思想，是什么原因导致了系统风险呢？经济变量方面的信息包括 GDP 高于预期值的政府公告或中央银行的突然加息等。这些整体经济范围内的因子 F（脚标为 j）对不同证券有着不同的影响，这可以用下式中 b_{ij} 或 β 的不同值来反映：

$$\begin{aligned} m_t &= \sum_j b_{it}\ (F_j - EF_j)_t \\ &= b_{11}\ (F_{1t} - EF_{1t})\ + b_{12}\ (F_{2t} - EF_{2t})\ + \cdots \end{aligned} \tag{27}$$

其中，期望算子以 $t-1$ 时刻及之前的信息为条件。例如，对某个特殊的公司而言，如果与利率变化相对应的 β 值为 0.5，那么利率每超过期望水平 1%，证券 i 的收益就会增加 0.5 个百分点（比期望收益大）。APT 的一个关键假设是，不同证券的异质型风险或特殊风险 ε_i 不相关，$\mathrm{Cov}\ (\varepsilon_i,\ \varepsilon_j)\ = 0$。

投资组合的收益

为了简单起见，假设在投资组合中有一个系统风险因子 F_t 和 n 种证券。根据定义，由 n 种证券组成的投资组合——证券的持有比例为 x_i——的收益为 R_t^p：

$$\begin{aligned} R_t^p &= \sum_{i=1}^n x_i R_{it} \\ &= \sum_{i=1}^n x_i\ (R_{it}^e + b_i\ (F_i - EF_t)\ + \varepsilon_i) \\ &= \sum_{i=1}^n x_i R_{it}^e + \left(\sum_{i=1}^n b_i x_i\right)\ (F_t - EF_t)\ + \sum_{i=1}^n x_i \varepsilon_i \end{aligned} \tag{28}$$

投资组合的收益等于预期收益的加权平均值，加上投资组合的 β 值（乘以与因子 F 有关的“消息”），再加上异质型风险 ε_i 的加权平均值。如果不同证券的异质型风险无关，那么 ε_i 中一部分为正，一部分为负，但加权和接近于 0。事实上，随着证券数目的增加，（28）右边最后一项趋于 0——异质型风险是可以分散的。因此，投资组合的收益由单个证券的预期收益和系统风险的收益——用单一经济信息（$F_t - EF_t$）来表示——所组成。

更为一般的方法

APT 的漂亮之处在于，它不需要效用函数的任何假设，投资组合的均值和方差是投资者所考虑的仅有的两个因素。模型实际上是一种机制（几乎是一种算法），在无风险套利机会瞬间即逝思想的基础上，模型能导出证券（或证券组合）预期收益的表达式。不足为怪的是，APT 确实需要一些（任意的）假设。我们假设投资者具有相同的预期，任何证券的收益 R_{it} 都与 k 个因子 F_{it} 线性相关：

$$R_{it} = \alpha_i + \sum_{j=1}^k b_{ij} F_{jt} + \varepsilon_{it} \tag{29}$$

其中，b_{ij} 是已知的因子权重。对（29）求数学期望，并假设 $\mathrm{E}\varepsilon_{it} = 0$，于是

$$R_{it} = \mathrm{E}R_{it} + \sum_{j=1}^k b_{ij}\ (F_{jt} - \mathrm{E}F_{jt})\ + \varepsilon_{it} \tag{30}$$

方程（30）说明，尽管每种证券都会受到所有因子的影响，但特殊 F_j 的影响大小取决于 b_{ij} 的值，对每种证券而言，这种影响的大小也是不同的。这就是不同证券收益 R_{it} 之间协方差的来源。我们假设在（30）中能持续增加因子，直到预料不到的收益部分 ε_{it} 满足：

对所有时期和 $i \neq j$ 的所有 i 和 j，有

$$E\ (\varepsilon_i \varepsilon_j)\ = 0 \tag{31a}$$

对所有股票和因子（和所有 t），有

$$E\left[\varepsilon_i\left(F_j - EF_j\right)\right] = 0 \tag{31b}$$

方程（31a）意味着不同证券的非系统性风险（或特殊风险）不相关，方程（31b）却意味着特殊风险与因子 F 独立。注意，对所有证券而言，度量系统风险的因子 F 都是相同的。现在，我们来进行一次“实验”，投资者构造一个包含 n 种资产的零 β 组合，净投资为零。零 β 组合必须满足：

$$\sum_{i=1}^{n} x_i b_{ij} = 0 \quad j = 1, 2, \cdots, k \tag{32}$$

投资为零的假设意味着

$$\sum_{i=1}^{n} x_i = 0 \tag{33}$$

于是，由（33）可知，有些 x_i 小于零，即卖空股票并将收益投资于其他股票。讨论的下一个部分将引入套利内容。如果投资者不投入任何资金，且零 β 组合能获得非零期望收益，那么无风险套利利润就会产生。这是对投资组合预期收益所施加的套利条件，因此，使用（28）后我们有

$$\begin{aligned}
R_t^p &= \sum_{i=1}^{n} x_i R_{it} \\
&= \sum_{i=1}^{n} x_i\left(ER_{it} + \sum_{j=1}^{k} b_{ij}\left(F_{jt} - EF_{jt}\right) + \varepsilon_{it}\right) \\
&= \sum_{i=1}^{n} x_i ER_{it} + \left(\sum_{i=1}^{n} x_i b_{i1}\right)\left(F_{1t} - EF_{1t}\right) \\
&\quad + \left(\sum_{i=1}^{n} x_i b_{i2}\right)\left(F_{2t} - EF_{2t}\right) + \cdots + \sum_{i=1}^{n} x_i \varepsilon_i
\end{aligned} \tag{34}$$

使用（34）和（32），根据充分分散的投资组合右边最后一项接近于 0 的假设，我们有

$$R_t^p = \sum_{i=1}^{n} x_i ER_{it} \equiv ER_t^p \tag{35}$$

由定义可知第二个等式成立。因为任意构造的投资组合的实际收益 R_t^p 都与期望收益 ER_t^p 相等，所以这个收益的波动率为零，从而没有任何风险。于是，套利理论认为，无风险收益必须等于零：

$$ER_t^p = \sum_{i=1}^{n} x_i ER_{it} = 0 \tag{36}$$

第一次看上去似乎有些奇怪，无风险组合却不能获得无风险利率 r_t。这是因为，无风险资产是可行的投资资产，与其他资产一样，我们可以用它的多头和空头头寸来使净收益为零。

现在，我们必须使用基于线性代数的证明。在（31）、（32）、（33）和（36）给定的条件下——这就是众所周知的正交条件，可以证明，任意证券 i 的期望收益可以写成因子权重 b_{ij} 的线性组合。例如，对两因子模型而言：

$$ER_i = \lambda_0 + \lambda_1 b_{i1} + \lambda_2 b_{i2} \tag{37}$$

我们注意到，（30）中的 b_{i1} 和 b_{i2} 对于证券 i 而言有些特别。证券 i 的期望收益等于证券特定的 β 值乘以权重 λ_j，对所有证券而言权重 λ_j 都相同。因此，λ_j 可以解释为证券必要的超额期望收益，因为证券对第 j 个因子（如 GNP 或利率）敏感。

λ_j 的解释

就此时此刻而言，假设我们知道 b_{i1} 和 b_{i2} 的值。我们可以将 λ_j 进行如下解释。考虑零 β 组合（如 $b_{i1} = 0$ 且 $b_{i2} = 0$），期望收益为 ER_z。如果无风险借贷存在，那么 ER_z 等于无风险利率 r。使用（37），我们可以看出

$$\lambda_0 = ER_z \text{（或者 } r\text{）} \tag{38}$$

接下来考虑 $b_{i1} = 1$ 且 $b_{i2} = 0$ 的投资组合，期望收益为 ER_1，代入（37）后我们就得到

$$\lambda_1 = ER_1 - \lambda_0 = E\ (R_1 - R_z) \tag{39}$$

类似地

$$\lambda_2 = E\ (R_2 - R_z) \tag{40}$$

因此，(37) 的一种替代形式是

$$ER_i = ER_z + b_{i1}E\ (R_1 - R_z)\ + b_{i2}E\ (R_2 - R_z) \tag{41}$$

于是，APT 的一种解释是，证券 i 的期望收益取决于它对因子承载（loading）（即 b_{ij}）的敏感性。另外，每个因子承载（如 b_{i1}）都是超额期望收益 $E\ (R_i - R_z)$ 的权重，也就是说，如果投资组合的期望收益关于第 1 个因子的 β 值为 1，关于其他因子的 β 值为 0，那么 β 值为 1 的投资组合能模拟第 1 个因子 F_1 中预料不到的运动。

APT 的应用

APT 是一个因子模型，可以用两个方程来对其进行概括：

$$R_{it} = \alpha_i + \sum_{j=1}^{k} b_{ij}F_{jt} + \varepsilon_{it} \tag{42a}$$

$$ER_{it} = \lambda_0 + \sum_{j=1}^{k} b_{ij}\lambda_j \tag{42b}$$

其中，$\lambda_0 = r_t$ 或 ER_z。注意，在 (42b) 中，λ_0 和 λ_j 对所有资产（横截面上）而言都是常数。时间序列方程 (42a) 对于 R_{it} 而言不是一个预测方程，因为因子也是在 t 时刻来进行度量的，方程度量的是因子的同时性风险暴露。

如果无风险利率不存在，那么必须估计 λ_0 的值。当所有的 β 值为 0 时，因为 λ_0 是某个投资组合的期望收益，所以 λ_0 为零 β 利率。b_{ij} 可以解释为资产 i 在因子 j 上的风险暴露数量，λ_j 可以解释为风险暴露的“价格”。所以资产的期望收益溢价等于因子 F_j 的单位风险暴露 β_j 乘以 λ_j——这个 β 值的风险价格。决定平均收益的是风险承载或 β 值，而不是公司的其他特征。例如，“价值股”（即面值市值高的股票）能获得一个高的平均收益，不是因为它们有较高的面值市值比，而是因为它们在面值市值因子 F_j 上有较高的风险承载（即大的 β_{ij} 值）。

APT 的（典型）使用方法如下。R_{it} 对一组因子 F_{it}（如通货膨胀率、GDP 增长率、利率等）进行第一次时间序列回归，由此可以得到 α_i、b_{i1} 和 b_{i2} 等的估计值。对 $i = 1,\ 2,\ \cdots,\ m$ 的证券重复上述过程，由此可以得到每个 β 的 m 个估计值。在进行第二次回归时，改变 m 种证券的 b_i 值，于是 b_i 就是 (42b) 中右边的变量。因此，在方程 (42b) 中，b_{ij} 相对于 m 种证券而言是不同的，λ_j 对所有证券而言都是相同的，在横截面回归方程中，$\bar{R}_i$ 对 b_{ij} $(i = 1,\ 2,\ \cdots,\ m)$ 进行回归就可以将 λ_j 估计出来。无风险利率为常数，因此它是横截面回归中的常数项。

上述估计是一个两阶段程序，但是（原则上至少是这样）存在一种更好估计程序来同时对方程 (42a) 和 (42b) 进行估计，这就是著名的因子分析。因子分析选择一部分因子 F_j，使每个方程残差项之间的协方差（近似）为零（即 $E(\varepsilon_i\varepsilon_j) = 0$），这与投资组合充分分散的假设一致。当下一个因子的增加无法改进解释时，就要停止增加因子 F_j。因此我们同时对 F_j 的数目和相应的 b_{ij} 进行了估计。于是，可以用横截面估计方程 (42b) 来对 λ_j 进行估计。

然而，因子分析在解释结果时会出现一些问题。第一，b_{ij} 和 λ_j 的符号是任意的，可能出现相反的情形。（例如，正 b_{ij} 负 λ_j 与负 b_{ij} 正 λ_j 在统计上没有区别）。第二，尺度变换问题存在，因为如果 β_{ij} 翻倍 λ_j 减半，那么结果仍然成立。最后，如果使用不同的样本数据来反复进行回归，我们不能保证同一因子以相同的顺序出现。因此，APT 模型唯一的一个优先约束条件是一些 λ_j 和 b_{ij}（在统计上）不为零：根据这一结果是无法获得多少经济直觉的。

我们几乎没有讨论 APT 的检验，原因是尽管模型结构非常宽泛（基于套利观点和其他几个限制性很弱的假设），这个模型的应用还是很困难。与 b_{ij} 和 λ_j 的解释问题一样——我们不能事先“指定”符号（即要么正要么为负），我们也存在如下问题：在时间轴上，b_{ij} 和 λ_j 不是

常数。

CAPM 和 APT

到目前为止，读者已经清楚，这两种均衡期望收益模型是建立在非常不同（行为）假设基础上的。经常将 APT 称为多因子模型。根据这一术语，标准 CAPM 模型可以认为是 APT 非常特殊的一种情形，即 APT 的单因子版本，唯一因子就是市场投资组合的期望收益 ER_m。如果假设证券 i 的收益生成方程只取决于一个因子，将该因子取为市场投资组合的收益，那么由 APT 有

$$R_{it} = \alpha_i + b_i R_{mt} + \varepsilon_{it} \tag{43}$$

可以证明，单因子 APT 方程（43）意味着期望收益为，

$$ER_{it} = r_t + b_i\ (ER_{mt} - r_t) \tag{44}$$

这与 CAPM 的均衡收益方程一致。APT 模型涉及一些细微之处，但从直观的角度来对其进行解释并不是很容易的事情。它的主要思想是：

(ⅰ) 在一个零风险投资组合（即零 β 组合）构造的基础上，模型提供了确定均衡收益的一种结构，但不需要现金投资。

(ⅱ) 这些条件再加上如下假设——线性因子权重假设：为了使特殊风险充分地小，证券数目必须足够大——允许我们对期望收益方程的参数施加正交限制条件。这些限制条件导致期望收益方程取决于资产 i 的因子承载 b_{ij} 和权重 λ_j，对不同资产而言 λ_j 是相同的。

(ⅲ) APT 不依赖于效用函数的任何假设，投资者仅考虑投资组合的均值和方差。然而，APT 要求预期同质。

(ⅳ) 当我们考虑 APT 的实证应用时，APT 包含一些随意的选择（例如，哪些因子 F_j 是合适的?）。APT 的解释很困难（例如，对 b_{ij} 和 λ_j 的符号事先无法进行限制）。

7.5　小结

* 夏普比、Treynor 指数和 Jensen 指数（“Jensen α”）都是绩效度量指标，可以用来对各种投资组合或投资策略进行排序。Jensen 指标可以推广到包含更多因子——而不仅仅是市场收益——的情形。

* 尽管实证文献主要使用标准 CAPM 和零 βCAPM，还是可以对 CAPM 按照多种方式进行扩展。

* 单因子模型（SIM）是任意股票或股票组合收益的线性统计表示，模型假设，唯一的因子对资产收益会产生系统性影响，这个模型可以大大简化收益、最优组合权重的分析和在险价值的计算。

* APT 是基于无套利原理的多因子期望收益模型。一段时期内，收益受到多个因子的影响，因子的 β 值可以解释横截面的平均收益。

第8章

CAPM 和 APT 的经验证据

目的

- 说明如何使用时间序列和横截面方法来对 CAPM 模型（和零βCAPM 模型）进行检验。
- 证明多因子模型能很好地解释横截面的股票市场平均收益。股票收益的横截面部分取决于账面市值比和规模两个变量所对应的β，但不取决于 CAPM 模型中的β。
- 说明 CAPM 模型中的β只能解释债券（或国库券）与股票在横截面上的平均收益差异。
- 概述如何将 APT 视为多因子均衡模型。

8.1　CAPM：时间序列检验

在同质预期和无风险利率（不是随机的）存在的假设下，“标准” Sharpe－Lintner CAPM 模型的直接含义是均值－方差有效。最为直接的检验是对超额收益回归方程

$$ER_{it} - r_t = \alpha_i + \beta_i\ (ER_m - r)_t \tag{1}$$

中的 $\alpha_i = 0$ 进行检验。我们假设个体收益在短期内独立同分布，也考虑资产之间的同时相关性，$E(\varepsilon_{it}\varepsilon_{jt}) \neq 0$。于是，在收益服从联合正态分布的假设下，使用面板数据（N 种资产，$t = 1, 2, \cdots, T$）和极大似然估计法就可以对（1）进行估计。但是，因为所有方程的自变量都相同，所以（α，β）的极大似然估计等价于对每一个方程都进行 OLS 估计：OLS 估计的残差可以用来构造 N 种资产的（同期）协方差矩阵 $\Sigma = \hat{e}\hat{e}'/T$。但 Wald 检验和似然比检验的渐进结果在有限样本中会受到样本容量的影响（参见 Campbell、Lo 和 Mackinlay（1997），第 5 章）。然而，对小样本而言也存在对 $\alpha = 0$（$N \times 1$）适用的多种检验方法。准确的检验功效（在假设为真的条件下拒绝零假设的概率）是 T 的递增函数（正如我们所期望的那样），但对 N 的大小非常敏感，从而要取一个较小的值（小于 10）。在条件异方差或序列相关性存在时就可以使用 GMM，尽管我们对 $\alpha = 0$ 的小样本检验性质不是很清楚。在这种情形中，为了获得有效样本中 α 的经验分布，可以（对各种分布）使用蒙特卡罗模拟（MCS）方法或对残差使用自助法（bootstrap）。

Black 零βCAPM 模型的检验方法是在回归方程中使用实际收益：

$$R_t = \alpha + \beta R_{mt} + \epsilon_t \tag{2}$$

其中，R 是由 N 种资产的收益所构成的 $N \times 1$ 向量，（α，β）都是 $N \times 1$ 向量。零假设为

$$H_0\!: \alpha = (e - \beta)\ \gamma$$

e 是 $N \times 1$ 单位向量，$\gamma = \mathrm{E}\ (R_{0B})$ 是零β组合的期望收益，通常将其视为无法观测的参数。使用迭代系统估计量的面板数据回归方法就可以给出带约束条件的β和γ（和协方差矩阵Σ）的

估计值，因此，使用标准（渐进）的 Wald 检验和似然比检验是可行的。

20 世纪 70 年代的早期研究发现 $\alpha_i = 0$，这些研究倾向于支持 Sharpe – Lintner CAPM 模型，但后续研究（如 Campbell、Lo 和 Mackinlay（1997））发现这一结论并不成立，Cochrane（1996）直接对“条件 CAPM 模型”进行了估计，模型中超额市场收益的影响用红利 – 价格比或期限溢价来进行“调整”。因此，市场收益对资产（或资产组合）收益的影响取决于那些反映“商业周期状态”的变量。他发现，对根据规模进行分组的投资组合收益而言，定价误差（即 Jensen “α”）是标准（无条件）CAPM 模型定价误差的一半。

8.2　CAPM：横截面检验

CAPM 模型认为，横截面上的股票平均收益差异线性地（唯一地）依赖于资产的 β 值。（这就是证券市场线，SML。）对这一假设进行检验的第一个问题是，单只股票收益的波动是如此的剧烈以至于人们不能拒绝不同股票之间的平均收益相同的假设。就单只股票而言，$\sigma \approx 30\% \sim 80\%$（每年），因此，即使 T 很大，$\sigma/\sqrt{T}$也仍然很大。一种解决方法是对投资组合中的股票进行分组，分组目的是为了使平均收益之间的差异达到最大——如果不存在这种平均收益差异，就无法对 CAPM 模型进行检验。根据原理，任何分组方法都是可行的，但基于符号 $A-E$、$F-J$ 等的分组并不能产生一个性质良好的平均收益差异。然而，根据“规模”和“账面市值比”进行分组是较为常见的方法，由此可以得到一个性质较好的平均收益差异。（尽管早期的实证研究都是根据资产的 β 值来进行分组。）

第二个主要问题是 β 值存在度量误差。有多种方法可以使这一问题的影响达到最小。一种方法是为个体股票指定一个“投资组合 β 值”，“投资组合 β 值”的数目较小，使用少数几个投资组合（如 10 个左右）的收益时间序列就可以估计出这些“投资组合 β 值”——人们认为这种分组方法可以使 β 的估计误差达到最小。为了考虑一家公司在不同时刻有不同 β 值的情形，需要将上述方法进行扩展，并使用滚动回归技巧。后面将对这一问题进行详细讨论。

横截面检验采用的是两阶段程序（下面将对其进行更多的讨论）。在整个样本中 β_i 为常数的假设下，对每种资产进行第一阶段的时间序列回归：

$$R_{it} - r_t = \alpha_i + \beta_i\ (ER_m - r)_t + \varepsilon_{it} \tag{3}$$

每种证券的 β_i 估计值可以用来进行第二阶段的横截面回归。因此，k 种证券的月度样本平均收益 $\bar{R}_i$（通常是在整个样本内）对第一次回归中 $\hat{\beta}_i$ 值再进行回归：

$$\bar{R}_i = \lambda_0 + \lambda_1 \hat{\beta}_i + v_i \tag{4}$$

将（4）和标准 CAPM 模型的表达式进行比较

$$R_i = r + \beta_i\ (R_m - r)\ + \varepsilon_i \tag{5}$$

我们期望 $\lambda_0 = \bar{r}$，$\lambda_1 = \bar{R}_m - \bar{r} > 0$，$\bar{R}_m$ 和 $\bar{r}$ 上的横线表示样本平均值。在进行第二阶段回归时，在对 CAPM 模型进行更强的检验时要注意的是，只有 β_i（$i = 1, 2, \cdots, k$）才会影响 $\bar{R}_i$，方程（5）中的其他横截面变量（如收益的方差 $\sigma^2_{\varepsilon t}$）都不显著。

注意，在方程（4）中，对所有（横截面）资产而言 λ_0 和 λ_1 都是常数。如果无风险利率不存在，那么就必须对 λ_0 进行估计，λ_0 是所有 β 为 0 的投资组合的期望收益，也就是零 β 利率。

一个严重的计量经济问题是，在第一阶段的时间序列回归中，$\hat{\beta}_i$ 的估计值可能是无偏的，但它的度量存在误差。因此，在第二阶段的回归方程（4）中，我们发现了古典的“变量误差”（errors – in – variables）问题，这意味着 λ_1 的 OLS 系数会向下偏斜。如果 β_i 的真实值与证券误差的方差 $\sigma^2_{\varepsilon t}$正相关，那么证券误差的方差可以作为 β_i 真实值的一个替代，因此，如果 $\hat{\beta}_i$ 存在度量误差，那么 $\sigma^2_{\varepsilon t}$在第二阶段回归中可能显著。最后要注意是，如果误差项 ε_{it}的分布是

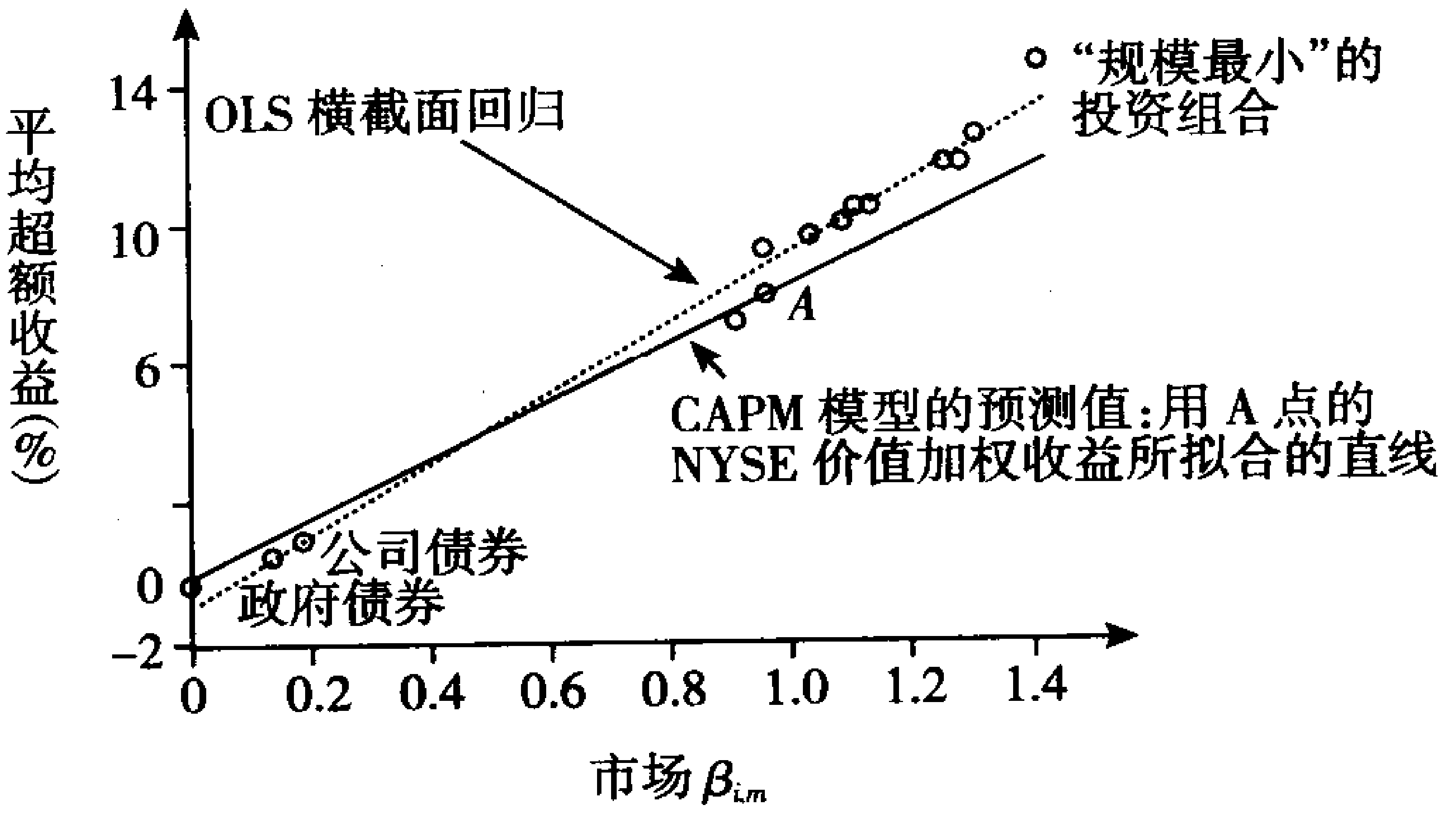
非正态分布（如正偏斜或肥尾），那么这将导致错误推断。特别地，横截面回归分析中的残差具有正偏斜特征，这意味着残差风险与资产收益之间存在某种联系，而在正确的模型中却不存在这种联系。

作为这些早期研究的一个说明，考虑 Black、Jensen 和 Scholes（1972），在第一阶段的时间序列回归中，他们使用的是1926—1966年的月度收益数据。根据单个证券 β 值的大小将所有证券分为10组，每一组就是一个投资组合（也就是说，用滚动的5年期估计窗口来对单个证券的 β_i 进行时间序列估计，然后根据 β_i 的大小来构造投资组合），他们用这种方法来使异方差问题和估计误差问题的影响达到最小。就这10个投资组合而言，在35年的时期上，将月度收益 R_{it}^p 对 R_{mt} 进行回归：

$$R_{it}^p = \hat{\alpha}_i + \hat{\beta}_i R_{mt} \tag{6}$$

这就得到了投资组合的 β 值，根据这10个 β 值，为每只股票“指定”一个 β 值。在第二阶段的横截面回归分析中，所有个体股票的月度平均超额收益 $\bar{R}_i - \bar{r}_i$ 对所“指定”的投资组合 β 值进行回归。CAPM 模型认为方程（4）中系数 λ_1 在统计上应该显著为正。

Cochrane（2001）对上述研究进行了更新，给出了 CAPM 模型的一个简单而又清楚的检验方法。首先，根据规模（即市值）将 NYSE 中所有股票分成10组，由此得到10个投资组合，同时考虑由公司债券所组成的投资组合和由政府债券所组成的投资组合（即总共有12个投资组合）。接下来，用时间序列回归方法分别估计12个投资组合的 β 值。最后，将每个投资组合的样本平均（$t=1, 2, \cdots, T$）收益 $\bar{R}_i$（$i=1, 2, \cdots, 12$）对12个 $\hat{\beta}_i$ 值进行回归。债券投资组合的 β 值较低，平均收益也较低。“根据规模分组的”投资组合收益与相应 β 值之间是一种正相关关系（图1）。所有结果似乎都支持 CAPM 模型。然而，这里存在两个问题。第一，如果我们画出 SML（$\bar{R}_i = \bar{r} + (\bar{R}_m - \bar{r})\beta_i$），那么投资组合收益的绝大多数点都位于 SML 之上，而不是位于 SML 直线中——这个问题也许不是很严重。第二，用 OLS 回归所得到的“最小规模”投资组合的平均收益位于 SML 之上（图1），这就是所谓的“小公司效应”（Banz（1981））——小公司（即低市值公司）所获得的平均收益超过经风险调整后的 SML 值。CAPM/SML 似乎能合理地解释债券和股票之间的相对收益，但不能对“根据规模进行分组”的投资组合在横截面上的平均收益进行合理解释。正如我们在下面将看到的那样，CAPM 模型的这种“成功”是非常脆弱的，当我们根据“账面市值比”来将股票分成10组时，β 值就无法解释这10个投资组合的收益。



按规模进行分组的、按照价值进行加权的10个投资组合（NYSE）：美国1947年以后的数据

Fama－MacBeth滚动回归

Fama－MacBeth（1973）给出了一种非常有用的回归方法，这种方法涉及“滚动”横截面回归。方法的本质在于，在每一个时期上分别进行横截面回归，由此得到所选择横截面变量（如β）的系数的时间序列，在此基础上进行各种检验。为了说明这一点，假设任意月份t内有N个企业，收益为R_t，这些收益的横截面β值是已知的。对CAPM模型进行检验的最简单方法是在任意月份的横截面上进行OLS回归：

$$R_t = \alpha_t e + \gamma_t \beta + \theta_t Z + \epsilon \tag{7}$$

其中，

$R_t =$（R_1，R_2，…，R_N）是t时刻的$N\times 1$月度横截面超额收益向量；

$\beta =$（β_1，β_2，…，β_N）是CAPM模型中的$N\times 1\beta$值向量；

$\gamma_t = t$时刻的横截面系数（β风险的“价格”）；

$\alpha_t = t$时刻的截距估计值；

$e =$元素全为1的$N\times 1$向量；

$Z =$其他横截面变量（如账面市值比）；

$\epsilon =$横截面上的$N\times 1$误差项向量。

对任意t而言，CAPM模型意味着，$\alpha_t = \theta_t = 0$且$\gamma_t > 0$。Fama－MacBeth程序是按照如下方式来对所有时期上的参数估计值进行平均。对月份$t = 1，2，\cdots，T$而言，重复回归（7）就得到α_t、γ_t和θ_t的T个估计值。然后，对这些参数的时间序列进行检验来考察$\alpha \equiv E(\alpha_t) = 0$、$\theta \equiv E(\theta_t)$和$\gamma \equiv E(\gamma_t) > 0$是否成立（即$\beta$是否存在正的风险溢价）。如果在一段时期内收益独立同分布且服从正态分布，那么下述统计量服从学生t分布（自由度为$T-1$）

$$t_\gamma = \tilde{\gamma}/\tilde{\sigma} \tag{8}$$

其中

$$\tilde{\gamma} = \sum_{t=1}^{T} \hat{\gamma}_t / T,$$

$$\tilde{\sigma}^2 = \frac{1}{T(T-1)} \sum_{t=1}^{T} (\hat{\gamma}_t - \hat{\gamma})^2$$

使用类似程序就可以得到t_α和t_θ的值，因此，我们很容易对CAPM模型的限制条件进行检验。上述程序的一个问题是，“实际的”β值未知，这就导致了“变量误差”（error－in－variables）问题。一种方法是对β使用工具变量，然而，为了使这一问题的影响达到最小，最为普遍的做法是，将单个收益进行分组，（比方说）分成100组（例如，根据规模或账面市值比来进行分组），计算这100个投资组合的β值，然后从这100个β值中为每种证券指定一个β值（参见附录）。

Fama－MacBeth（1974）使用上述程序，针对1935—1968年期间的每个月来对横截面投资组合收益方程分别进行估计，当然也将β_i^2和$\sigma_{\varepsilon i}^2$（由第一阶段的回归得出）包含在方程中：

$$\bar{R}_i^p = \lambda_0 + \lambda_1 \beta_i + \lambda_2 \beta_i^2 + \lambda_3 \sigma_{\varepsilon i}^2 + \eta_i \tag{9}$$

他们发现，λ_2和λ_3的（时间序列）平均值在统计上显著不为0，$\lambda_1 > 0$，得到的结果支持标准的CAPM模型。

在实践中，可以用CAPM模型来估计公司的权益成本，权益成本又可以用来计算资本加权成本（WACC），资本加权称本是投资项目评估中所使用的折现率（参见Cuthbertson和Nitzsche（2001*a*））。如果公司的估计β值向上偏斜，那么这会使CAPM模型中的必要收益率（hurdle rate）变得太高，从而使公司放弃一部分能赢利的投资项目（即内部收益率不超过向上偏斜的必要收益率）。有多种方法可以用来计算β的值（Blume（1975）），但是Bartholdy和Peare（2003）指出，如果我们使用某个市场指数（如S&P500）来估计β的值，而超额平均市场收益是用某个不同的市场指数（如NYSE指数）来进行计算的，那么权益融资成本的估计值

就是有偏的。然后，他们说明了如何使用 Fama - MacBeth 程序来得到“实际”权益融资成本的无偏估计值，即使所使用的是“市场指数”的某个替代指标。首先，在方程（3）中使用市场指数（如S&P500）的60个（比方说）月度数据（直到时间t为止）来得到每家公司β_i^{pr}的估计值。然后，使用下一个月度数据（$t+1$）来对（7）进行横截面回归，由此得到γ_t的第一个估计值。向前滚动回归，然后计算$\tilde{\gamma}$的平均值。他们认为权益资本的成本为

$$\begin{aligned} E_t R_{i,t+1} - r_t &= \beta_i E_t \ (R_{m,t+1} - r_t) \\ &= \beta_i^{pr} \ [E \ (R_m^{pr} - r) \ / \rho_{pr,m}^2] \\ &= \beta_i^{pr} \tilde{\gamma} \end{aligned}$$

$\beta_i^{pr}\tilde{\gamma}$就是它的无偏估计值。因此，在上述程序中，我们可以使用市场投资组合的任何指标（如S&P500，NYSE 指数）来获得权益资本成本的无偏估计值。当然，方法的使用仍然取决于 CAPM 模型是否是正确的期望收益模型，如果 CAPM 模型不能很好地“拟合数据”，那么即使获得了资本成本的无偏估计值，较大的标准差也不会让人感到满意。

Roll 批评

Roll（1977）表明，对事后有效的任何投资组合而言（记为q），平均收益与β值在样本数据中存在精确线性关系。因此，CAPM 模型的唯一检验含义是：市场投资组合是均值 - 方差有效组合。如果市场投资组合是均值 - 方差有效的，那么 CAPM/SML 在样本中必须成立。因此，实证检验中 SML 被违背的事实可能意味着，研究人员所选择的投资组合不是真正的“市场投资组合”。除非研究人员非常自信他所选择的投资组合就是市场投资组合（可能包括土地、商品、人力资本、股票和债券），否则基于 SML 的检验就是多余的，对 CAPM 模型的证实也不会提供任何帮助。尽管如此，持批评态度的研究人员还在继续考察 CAPM 模型的实证有效性，即使他们所使用的市场投资组合指标是错误的。这是因为，人们仍然感兴趣的是，一个特殊的实证模型——即使不完美——在多大程度上能解释均衡收益。我们始终可以考察，对市场投资组合的各种选择而言，第二阶段的回归结果是否稳健。

8.3 CAPM、多因子模型和 APT

APT 的一种直接检验方法是使用某种形式的因子分析法。Roll 和 Ross（1984）对1962—1972 年期间每日数据的42组股票进行因子分析，每组包含30只股票。在第一阶段的回归分析中，他们发现，对绝大多数组别而言，大约5个“因子”就可以对R_{it}进行很好的统计解释。在第二阶段的回归分析中，他们发现，3个因子就足够了。然而，Dhrymes、Friend 和 Gultekin（1984）表明，用因子分析法来解释结果时存在一个问题，如果因子分析中所考察的证券数目增加，那么统计上显著的因子数目也将增加。另外，“因子”是经济变量的线性组合，这一结果也无法得到解释。因此，过去20年研究的中心是，在多因子回归分析法的基础上来解释横截面上的平均收益。

正如我们已经看到的那样，更多的最新研究都试图采用分组的方法来减小β的度量误差，并使所选择的投资组合在横截面上具有更大的平均收益差异。事实上，从 Fama 和 French（1993）的经典论文开始，注意力已经转移到多因子模型上。为了使横截面上的收益差异达到最大，我们要对股票进行分组，由此得到多个投资组合。频繁使用的分组方法是根据“账面市值比”（即“价值”）和市场价值（即“规模”）分别将股票分为5组，由此得到25个投资组合和横截面上的25个平均收益。投资组合β值的使用可以降低度量误差，也可以缓解如下问题的严重性：因为财务杠杆、公司规模、商业风险等因素的变化，个体股票的β值也可能随着时间的演变而发生变化。

美国：横截面数据

Fama 和 French（1993）发现，三因子模型能解释根据“规模和价值”对美国股票进行分类的 25 个投资组合的月度收益序列，三个因子分别是市场投资组合的收益、“根据规模进行分组的”投资组合的收益（即规模小的投资组合的收益减去规模大的投资组合的收益，R_{SMB}（small minus big portfolio））、“高账面市值比投资组合的收益减去低账面市值比投资组合的收益”（R_{HML}（high minus low（HML）book - to - market portfolio））。时间序列回归方程和横截面回归方程分别为：

$$R_{it} = \beta_{mi} R_{mt} + \beta_{SMB,i} R_{SMB,t} + \beta_{HML,i} + R_{HML,t} \tag{10}$$

$$\bar{R}_i = \lambda_m \beta_{mi} + \lambda_{SMB} \beta_{SMB,i} + \lambda_{HML,i} + \beta_{HML,i} \tag{11}$$

依次对 25 个投资组合进行时间序列回归，我们就得到了方程（10）中 25 个投资组合的 75 个 β 估计值。在方程（11）中，可以用这些横截面上的 β 值和 25 个月度平均收益 $\bar{R}_i$（比方说，240 个月的平均，假定数据的噪声给定，参见 Cochrane（2001），第 15 章）来估计 λ_s。Fama 和 French（1993）发现，对根据“规模和价值”进行分组的 25 个投资组合而言，投资组合的市场 β_{mi} 值云集在 0.8 ~ 1.5 的范围内，25 个月度平均收益之间呈现出较大的差异，在 0.25 ~ 1 的范围内变化（图2）。如果 CAPM 模型正确，那么平均收益和投资组合的市场 β 值就会完全正相关，图 2 中所有点都位于 45°线附近。因此，CAPM 模型的 β 值几乎不能解释股票平均收益 $\bar{R}_i$ 在横截面上的变化（正如我们已经注意到的那样，尽管它确实有助于解释股票与国库券（或公司债券）在横截面上的相对平均收益差异）。正如我们在下面将看到的那样，平均股票收益在横截面上的绝大部分变化都可以用 SMB 和 HML——“因子模拟”组合——的 β 值来进行解释，但不能用 CAPM 模型的 β 值来进行解释。

注意，如果将不同“规模”的收益（在某个给定的“账面市值比”分组中）连接起来，那么“规模收益”与市场 β_{mi} 之间将再次出现正相关关系（为了清楚起见，在图 2 中，我们只对 5 个“账面市值比”（BMV）分组进行了这样的操作，但这种正相关关系同样适用于另外四个“账面市值比”分组）。因此，小公司的高平均收益可以部分地由高市场 β_{mi} 值来解释。但是，根据 BMV 分组的结果则不支持 CAPM 模型。在某个固定的“规模”分组中，如果根据“账面市值比”来对 5 个子组的收益进行考察，那么平均收益与市场 β_{mi} 是一种负相关关系（参见图 2，这种负相关关系同样适用于另外四个“规模”分组——图 2 没有将这些点连接起来）。因此，股票的分组方式对 CAPM 模型的有效性会产生影响。但是，无论按照哪种标准来进行分组，CAPM 模型都应该成立，因为无论是个体的（平均）收益还是投资组合的（平均）收益，它们与市场 β 值之间的关系都应该是线性关系。因此，根据“账面市值比”分类的投资组合收益不能由 CAPM 模型中的 β 值来进行解释，这实际上是一种异像（相对于 CAPM 模型而言）。

图 3 表明，Fama - French 的三因子模型是成功的，图中画出了 25 个根据“规模和价值”进行分组的投资组合的预测收益和实际平均收益，与图 2 相比，它们更接近于 45°线。图 3 右上角用虚线连接起来 5 个点表示在“账面市值比”最大的分组内（即“低价型”或“价值型”）规模不同的 5 个投资组合，Fama - French 模型的拟合效果较好。但图 3 左下脚用实线连接起来的 5 个点表示在“账面市值比”最小的分组内（即“高价型”或“成长型”）规模不同的 5 个投资组合中，Fama - French 模型的拟合效果很糟糕。从图 3 可以看出，除了“成长型”外，Fama - French 三因子模型的预测收益能很好地拟合 25 个投资组合的实际平均收益，与图 2 相比，这些点更加接近于 45°线。

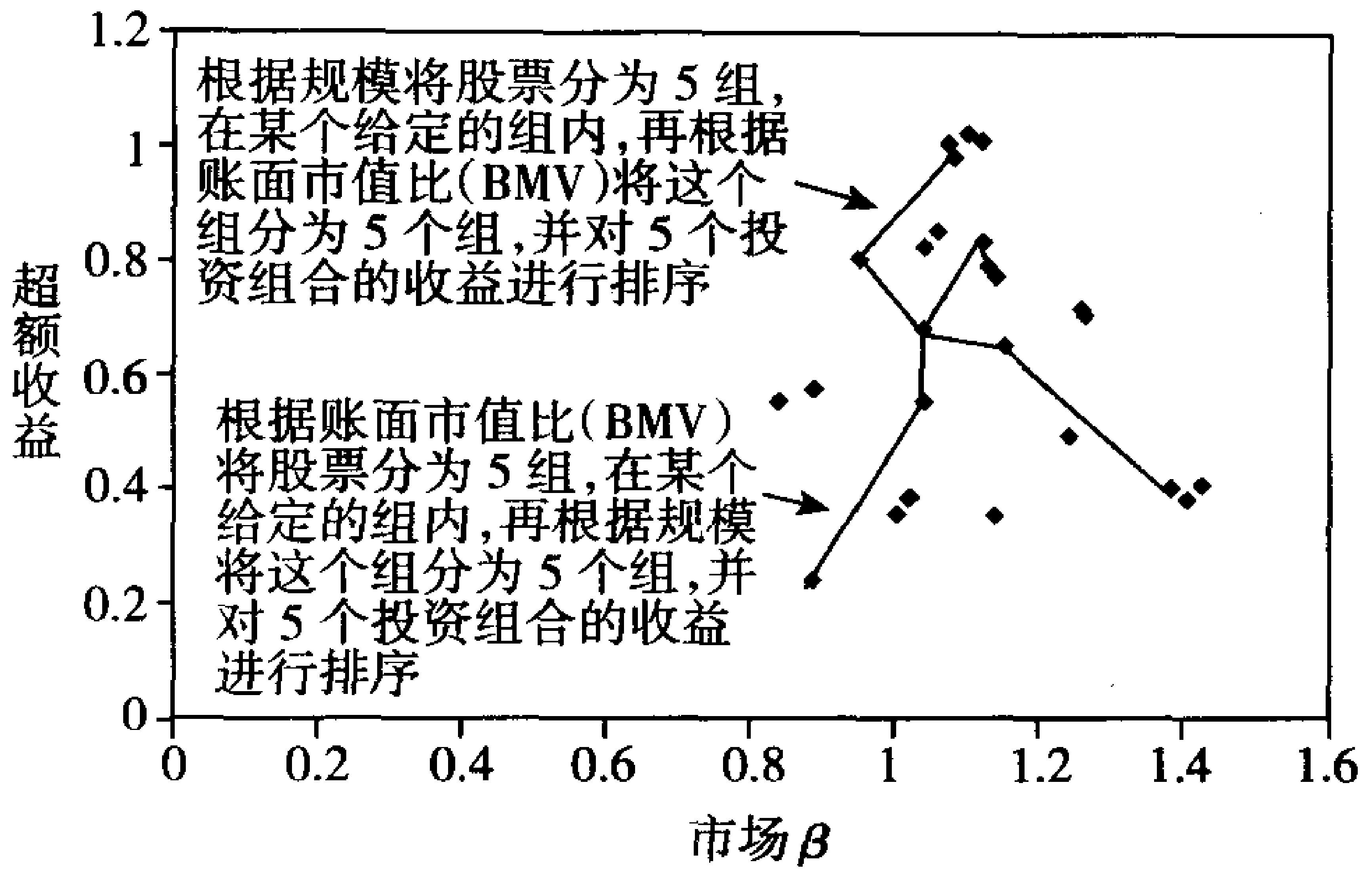
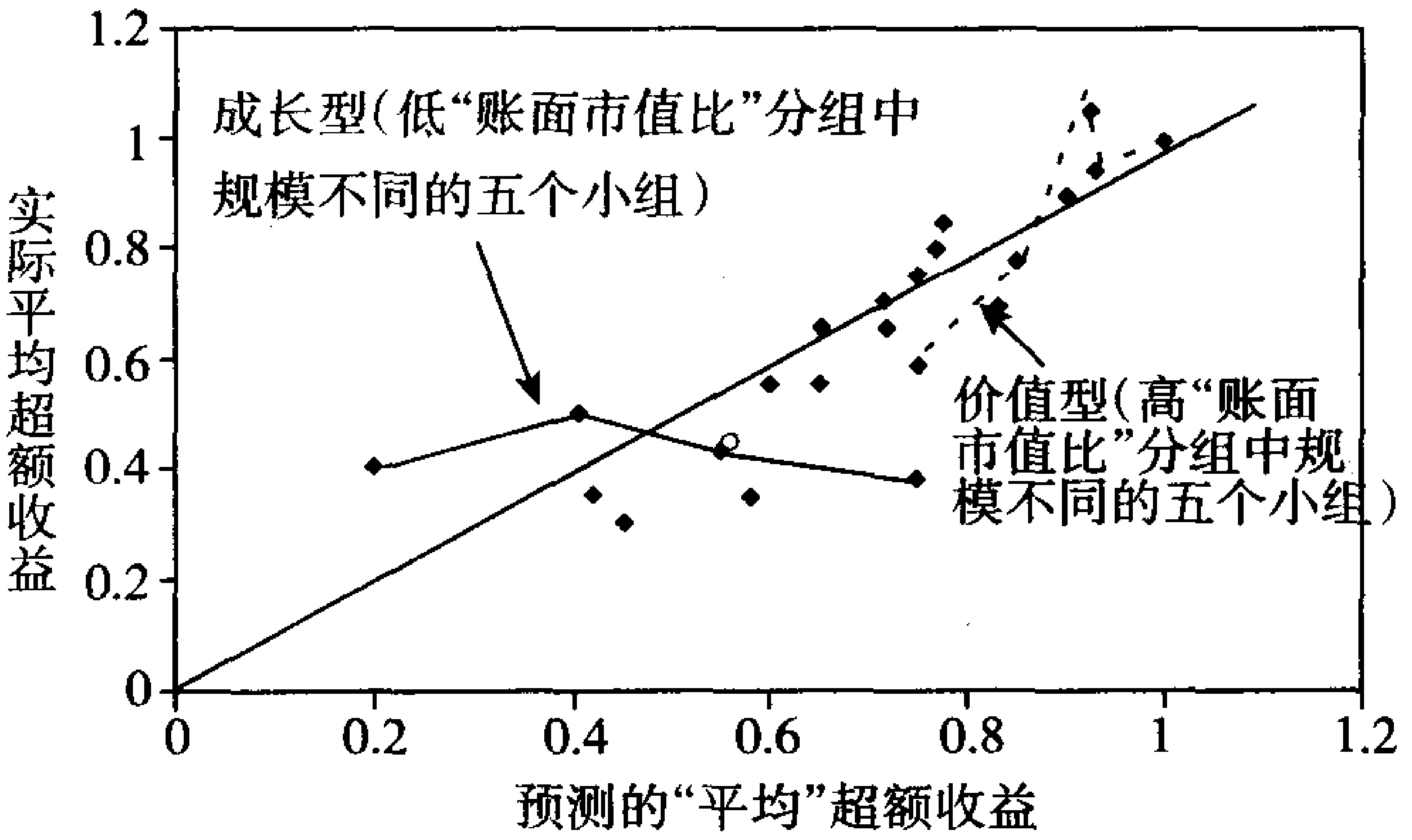


图 2　**平均超额收益与市场β（25 个根据“规模和 BMV”进行分组的投资组合）**



两条线连接的是某个面值市值比分组中规模不同的 5 个投资组合。我们仅给出了 BMV 最大和最小分组中的连线。如果将另外三个 BMV 分组中的连线给出来，那么所出现的结果（正相关关系）与价值型投资组合的结果相同，这说明，由 Fama - French 三因子模型所预测的收益可以用来对根据规模和 BMV 进行分组的投资组合平均收益进行预测。

图 3　**实际平均收益与用 Fama - French 三因子模型所预测的平均收益**

注意，在 20 世纪 80 年代中期，“规模效应”似乎消失了。Cochrane（2001）使用 1979—1998 年的数据研究表明，对根据“规模”进行分组的 10 个投资组合而言，月度平均收益云集在图 1 中的 A 点附近。因为这些点随机地位于 SML 附近，所以小公司溢价现象消失，10 个投资组合的市场β值都接近于 1.0，从而根据规模分组的投资组合收益与市场β值在 1979 年以后就不存在正相关关系，也就是说，文献中的异常“规模效应”消失（Banz（1981））。

我们仍然将 Fama - French 模型视为 APT 模型。如果在 25 个点的回归分析中 R^2 为 100%，那么三因子模型能完全拟合 25 个收益，从而不存在任何套利机会。Fama 和 French（1993）发现回归分析中 R^2 的大于 90%，因此 APT 是一个备选模型。

Fama 和 French（1996）将早期分析进行了扩展，发现模拟投资组合的 HML 和 SMB 因子能解释横截面上的平均收益，但股票是根据“新”标准来进行分类的，即价格乘子（价格收益比）和五年的销售增长率。但这些新投资组合的平均收益无法由 CAPM 模型的市场β来进行解释。然而，Fama - French 模型存在一个众所周知的“缺陷”，如果根据股票最近的收益绩效（即“动量股票”）来对股票进行分组，那么 HML 和 SMB 无法解释股票的平均收益，具体分析参见第 18 章。

我们仍然不清楚的是，哪些经济原因导致了 HML 账面市值比因子能获得定价，这个领域已经有了一些进展。Fama 和 French（1995）发现，典型的“价值型公司”（价格（从而账面市值比）较低）通常是陷入困境或濒临破产的公司。因此，HML 为商业周期的总量风险提供了一些信息，HML 投资组合的收益确实有助于解释 GDP 运动的证据也较为微弱（Liew 和 Vassalon（1999），Heaton 和 Lucas（1996））。小公司具有预示财务困境（如高杠杆比率）的特征，这一事实是真的，但也要注意的是，这种风险应视为一个风险因子（因为个体公司的特殊风险能以近乎零成本的方式来进行分散）。小公司效应也是一种“低价”现象，因为规模的其他度量指标如面值或员工数目等都不能预测横截面收益（Berk（1997））。最后要注意的是，那些确实有助于预测股票收益的宏观经济因子，如红利（收益）价格比、期限利差、违约风险利差等，也有助于预测经济衰退，这说明，使用因子模拟组合的 Fama - French 三因子模型背后隐含的是“衰退型”解释。

英国：横截面数据

Miles 和 Timmermann（1996）使用 Fama - MacBeth 程序对英国 1979—1991 年（12 年）的月度数据进行了考察。在每个月中，457 家公司的横截面回归方程为：

$$R_i = \lambda_0 + \lambda_1 \beta_i + \lambda_2 BMV_i + \lambda_3 \ln MV_i + \varepsilon_i \qquad (12)$$

其中，R_i 是月度收益，账面市值比 BMV_i 和市场价值（“规模”）ln（MV_i）两个解释变量都要滞后一年。在 12 年的时期中（1979—1991）就每一个月对这个回归方程重复进行回归，因此，横截面回归分析给出了每个横截面参数 λ_i 的 144 个估计值。他们发现，在解释平均收益（λ_2 的平均值为 0.35，t 统计量的平均值为 2.95）和正相关关系（也就是说，与高 BMV 相伴随的是高平均收益）时，BMV 是一个非常重要的横截面变量。他们也发现，CAPM 的 β 值、红利收益的 β 值、P/E 比率或杠杆比率都不会对横截面收益产生影响，只有最小规模投资组合存在“规模效应”，但“规模效应”的证据非常微弱。

Miles 和 Timmermann（1996）也报告了“因子”对平均收益影响的非参数自助检验（bootstrap test）结果，这种检验方法可以避免对回归检验所施加的线性性和正态分布假设。在 12 年中的每一个四月内，随机地将 457 只股票分成 10 个等权重的投资组合，然后计算 144 个月中每个投资组合的月度平均收益，最后对这些收益进行排序就得到最大平均收益、最小平均收益和平均收益差（等于最大平均收益减去最小平均收益）。重复程序 5000 次就得到“最大平均收益”、“最小平均收益”和“平均收益差”三个变量的分布，以及“随机收益”零假设下 1%、5% 和 10% 显著性水平下的临界值。当我们根据 BMV 和所计算的“最大平均收益”、“最小平均收益”和“平均收益差”来对投资组合的收益进行分组时，这些变量在 1% 的显著性水平下显著，但这对其他解释变量不成立。

例如，在投资组合随机排序的自助法检验中，最小（最大）收益 1% 显著性水平下的临界值为每月 0.83%（1.50%）。如果根据 BMV 进行分组，那么账面市值比最小（最高）的投资组合的平均收益为每月 0.75%（1.54%），这表明，根据 BMV 进行分组所得到的 BMV 与平均收益之间的正相关关系（不要求为线性）在 1% 的显著性水平下显著。如果根据其他“关键变量”进行分组，那么“规模”最小的一个投资组合的平均收益比用自助法得到的临界值低 1%，这意味着只有小公司才具有“规模效应”。因此，由非参数自助法所得到的结论与由回归方法所得到的结论一致，即 BMV 是横截面上平均收益的关键决定因素。

可供选择的“风险”因子

各种宏观经济变量（如通货膨胀、投资增长、劳动收入、消费财富比）在 APT 的解释中也经常用作因子（如 Chen、Roll 和 Ross（1986），Cochrane（1996），Lettau 和 Ludvigson（2001b）），这些宏观经济因子通常也能获得定价（也就是说，在横截面回归方程（11）中，在统计上是显著的），因此多因子 APT 模型成立。但是，这些宏观经济变量一般不能解释根据

规模和价值进行分组的横截面收益，这与 Fama - French 模型中的 SMB 和 HML 因子一样。Lettau和 Ludvigson（2001a）模型是一个例外，他们模型中的宏观经济变量是$z_t=(c/w)_t\Delta c_{t+1}$。对消费增长的任意水平而言，因子取决于衰退变量消费财富比 c/w，c/w 表明偏好是时变风险规避类型。明天消费的边际效用在这里不仅取决于消费增长率（与标准 CCAPM 模型一样），还取决于你是否处于衰退之中，即 c/w 变量。z_t 因子的 β 与 Fama - French 三因子模型中的 SMB 和 HML 因子的 β 一样，能够解释根据“规模 - 价值”进行分组的 25 个投资组合的横截面收益。

显然，上述实证工作表明，两个或多个因子在股票横截面收益中起着重要的决定作用，就目前而言，Fama 和 French（1993，1996）的工作在解释根据“价值”和“规模”进行分组的横截面收益方面处于“市场领导者”的地位。

交叉方程限制与 APT

在 Fama 和 MacBeth（1973）最早的论文中，第二阶段回归涉及 2000 只股票，从 100 个“投资组合”的 β 值中为每只股票指定一个 β 值，但“规模”和 BMV 等横截面变量有 2000 个观测值。因此，如果“规模”和“价值”在横截面上有较大的变化，那么根据有限个投资组合 β_i 值所得到的结果将是有偏的。Fama 和 French（1993）的研究使用的是 OLS，所隐含的假定是，（在任意时刻 t）横截面上异质性风险（即误差项 ε_{it} 和 ε_{jt}）之间的相关系数为零。APT 意味着，Fama - MacBeth 和 Fama - French“单方程”研究中没有考察时间序列和横截面参数之间的交叉方程限制。APT 可以表示为：

$$R_{it}=E(R_{it})+\sum_{j=1}^{k}b_{ij}F_{jt}+\varepsilon_{it} \tag{13}$$

R_i 为资产收益，F_j 为因子。期望收益可以表示为：

$$E(R_{it})=\lambda_0+\lambda_1 b_{i1}+\cdots+\lambda_k b_{ik} \tag{14}$$

因此

$$R_{it}=\lambda_0+\sum_{j=1}^{k}b_{ij}\lambda_j+\sum_{j=1}^{k}b_{ij}F_{jt}+\varepsilon_{it} \tag{15}$$

常数项 α_i 无约束的回归方程是（设 $\lambda_0=r$）：

$$(R_{it}-r_t)=\alpha_i+\sum b_{ij}F_{jt}+\varepsilon_{it} \tag{16}$$

比较方程（15）和（16）就可以看出，如果 APT 是正确的模型，那么方程（16）存在非线性交叉方程限制 $\alpha_i=\sum_{j=1}^{k}b_{ij}\lambda_j$。在时间序列横截面回归分析中，我们使用 NLSUR 方法，这种方法考虑了误差项之间的同期协方差（给出的估计值也更有效）。因此，方程（15）需要对 λ_j 和 b_{ij} 同时进行估计。

McElroy、Burmeister 和 Wall（1985）根据美国的投资组合收益发现，APT 的限制条件不成立。Clare 等用英国数据对方程（15）进行了估计，他们对 APT 的参数施加了限制，考虑了误差项之间的同期交叉相关性，结果发现，CAPM 模型的 β 风险价格（即市场收益 β 的系数 λ_β）为正，且在统计上显著。与之相对比的是，当误差项的方差协方差矩阵限制为对角矩阵时（即与 Fama - MacBeth 程序相似），λ_β 在统计上不显著。使用 NLSUR 所估计的 β 风险价格 λ_β 非常稳定，其他横截面变量，如“规模”、账面市值比和价格收益比等的 β 值，无法使解释能力增强。我们还不完全清楚的是，为什么对英国数据使用 NLSUR 方法所得到的结果与对美国数据使用 Fama - MacBeth 两阶段方法所得到的结果是如此的不同，而 NLSUR 方法确实能改进效率，但无法对任何偏差进行纠正（两种方法都存在“变量误差”偏差问题）。也许是在 APT 上所施加的限制导致了这种差异或英美两国的收益出现不同行为。因此，似乎出现的结果是，如果 CAPM 模型在英国成立，那么在美国就不成立。这一矛盾结果至今仍然是一个谜。

总之，对一个可接受的经验 APT 进行检验和考察的关键问题是，因子集合 F_{jt} 和所得到的 λ_j 值在不同的样本期上和不同的投资组合（如根据“规模”和“面值”进行分组）中是否为

常数。如果 λ_j 在不同的样本期上不同，那么因子 j 的风险价格就具有时变性（与理论相反）。尽管对 APT 的估计和检验已经有了很大的进展，上述问题的经验证据仍然不足。

蒙特卡罗模拟和自助法

在前面的章节中，在 CAPM 和 APT 模型检验中所使用的计量经济技巧存在很大的迷惑性，然而不幸的是，由于篇幅的限制，我们无法对这些重要问题进行深入讨论——这也是为什么有那么多计量经济文章的原因，一些文章指出计量经济方法在实际应用中确实存在局限性。在绝大多数计量经济文献中，假设 *niid* 误差的极大似然估计法和正确的模型设定（如线性）通常“位于树的顶端”，因为所给出结果的是渐进有效估计结果（OLS 和 GLS 是 ML 的特例）。然而，对序列相关性和异方差而言（如 White（1980），Newey - West（1987）），第一阶段的 GMM（对误差的协方差矩阵进行修正）不需要对误差项进行特殊的参数化处理，也不需要正态分布假设。在资产定价文献中，GMM 是更为常用的方法，因为它对误设模型的估计结果更为稳健。许多分布结果（如零假设下 χ^2 的检验统计量）只具有渐进性质，精确统计量需要正态分布假设，因为这个原因，所以蒙特卡罗模拟（Monte Carlo simulation）和自助法（bootstrap）的使用与日俱增。（Cochrane（2001）对资产定价模型中所使用的估计程序进行了最为漂亮、最容易理解的阐述，参见第 15 章和第 16 章）。

例如，在 CAPM 模型的检验中，为了对各种估计技巧进行评价，我们可能按照如下程序进行操作。CAPM 模型意味着

$$R^e_{it} = \alpha_i + \beta_i R^e_{m,t} + \varepsilon_{i,t}$$

其中

$$R^e_{it} = R_{i,t} - r_t$$

$$R^e_{m,t} = R_{m,t} - r_t$$

如果 CAPM 模型是正确的，那么我们期望 $\alpha_i = 0$。假设你根据 T 个月度观测值使用 OLS 对 10 个投资组合进行了估计，因此你就得到了 10 个投资组合的 α_i、β_i 和残差 $e_{i,t}$ 的估计值，也得到了样本的（同期）协方差矩阵 Σ。现在，在误差项为 *niid* 的假设下，使用样本协方差矩阵 Σ 就能生成拟合数据。接下来，在 CAPM 模型 $R^e_{it} = 0 + \beta_i R^e_{m,t} + \varepsilon_{i,t}$ 是正确模型的零假设下（β_i 为 OLS 估计值），我们来生成 10 个投资组合的收益（长度为 T）。我们假设 R^e_m 服从正态分布，T 个值从 $R^e_{mt} \sim niid\ (\bar{R}^e_m,\ \sigma\ (R^e_m))$ 中进行抽取，样本估计值与误差项独立。这些都是标准 ML 的假设。根据这些拟合数据，我们就可以使用各种计量方法（如 ML、GLS、OLS、一阶段或二阶段 GMM）来对 CAPM 模型 $R^e_{it} = \alpha_i + \beta_i R^e_{mt} + \varepsilon_{it}$ 进行估计，并对 $\alpha_i = 0$ 进行检验。将以上步骤重复 10 000 次（比方说），获得 10 000 个参数估计值，获得的 α_i 10 000 个检验统计量值。如果计量经济方法是“好”的方法，那么我们期望在 5% 的时间内（这就是检验的“规模”）在 5% 的显著性水平下拒绝 $\alpha_i = 0$ 的零假设。因为我们已经生成了 10 个投资组合拟合的时间序列数据（长度为 T），所以我们可以计算样本平均值，使用 OLS 或横截面 GLS（= 协方差矩阵为 $E\ (vv')$ 的 ML）或 Fama - MacBeth 滚动回归来对 $\bar{R}^e_i = \lambda_1 \hat{\beta}_i + v_i$ 进行检验。我们可以将模拟重复 10 000 次，检验 λ_1（>0）是否等于超额市场收益。

我们可以重复上述步骤，但不是从服从正态分布的误差项中进行抽取，而是从初始的 OLS 残差 e_{it} 和 $R^e_{m,t}$ 的原始样本数据中进行（随机）抽取，从而生成拟合数据序列 $R^e_{it} = 0 + \beta_i R^e_{m,t} + e_{i,t}$。此外，如果我们认为残差序列相关，那么我们就可以使用分块自助法（block - bootstrap）（如在长度为 3（比方说）的块中抽取原始数据 $\{R^e_{m,t},\ e_{it}\}$ 来捕获 MA（2）误差。用自助法得到的分布和用各种估计技巧对 $\alpha_i = 0$ 的检验反映了实际数据集合中市场收益和残差可能的序列相关性、异方差性、非正态性和非独立性等。

Cochrane（2001）对根据“规模”进行分组的 10 个 NYSE 投资组合分别在两个数据集合上进行了上述操作：二战后的 T（$= 876$）个月度数据，T（$= 240$）个短期月度数据。他发

现，对横截面回归而言，各种检验技巧的结果基本相同。现在考虑对 $\alpha_i = 0$ 的时间序列进行检验。MCS 假设残差为 *niid*，尽管 GMM（滞后 3 期）能对移动平均误差进行调整，零假设的拒绝频率还是与 ML 差不多，两者都是在 6% ~7% 而不是的显著性水平下决绝零假设。（众所周知，如果 GMM 使用长期滞后的不加权的谱密度矩阵，那么 GMM 拒绝零假设的频率要高得多。如果滞后长度为 24，那么 GMM 在这里的拒绝概率大约为 25% ~40%。）然而，使用实际数据的自助法检验结果表明，当残差不服从正态分布时，这种技巧是非常有用的。$ML\chi^2$ 检验的功效接近于 50%，如果显著性水平为 5%，那么检验的拒绝概率为 2.8%，而第一阶段的 GMM 估计（对异方差进行了纠正）能对“规模扭曲的 GMM”进行纠正，从而估计结果更加稳健。下一章将深入考察一些例子，说明 MCS 和自助法在对各种检验统计量的有限样本性质进行考察时的作用。所有模型都不是完全正确的，因此我们应该对各种模型的相对检验效率进行比较（如假设的合理性、内部一致性和简洁性），在使用各种技巧和检验方法时，也应比较各个模型对数据的拟合能力（参见 Inter alia、Cochrane（2001）和 Hendry（1995）对计量经济检验及其与经济理论之间关系的很有意思的讨论，这是一个非常有争议的话题，在（近似）相同条件下我们不能“重复实验”，但自然科学家却可以这样做）。

8.4 小结

* CAPM 模型的 β 值能解释股票和债券平均收益（横截面）之间的差异，但无法解释不同股票投资组合间的平均收益差异。

* Fama - French 的“高账面市值比投资组合收益减低账面市值比投资组合收益”因子和“规模因子”在很大程度上能解释根据“规模”和“账面市值比”进行分组的股票投资组合在横截面上的平均收益差异（对美国和英国而言）。这些“因子”可以松散地解释为，投资者为了对由衰退所导致的风险进行补偿，他们会要求一个较高的平均收益。这些因子代表的是“衰退溢价”（distress premia）。

* 如果根据最新绩效（如动量投资组合）来对投资组合进行分类，那么 Fama - French 三因子模型不能解释投资组合的横截面收益。

附录：Fama - MacBeth 两阶段回归

Fama - MacBeth 横截面回归方法需要股票 β 的估计值。CAPM 模型可以表示为：

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it} \qquad (\mathrm{A1})$$

其中，R_i 和 R_m 分别表示个体股票和市场投资组合的超额收益。Fama 和 MacBeth（1973）对 2000 只股票进行了研究，但他们只计算了 100 个投资组合的 β 值，然后从这 100 个 β 值中为每一个每只股票指定一个 β 值。程序如下：

(i) 使用 $t = 1 \sim 60$ 的月度观察值依次对每只股票估计方程（A1）中的 β_i 值。

(ii) 在 $t = 61$ 处，根据市场价值（“规模”）来构造 10 个组合，然后根据所估计的个体股票 β_i 值将每个组合分为 10 个子组合。在 $t = 61 \sim 72$ 期间（即一年内）计算 100 个投资组合的平均月度收益。现在，我们就有根据规模和 β 进行分类的 100 个平均收益 $\bar{R}_p$（$p = 1, 2, \cdots, 100$）。

(iii) 针对每一年来重复上述程序，由此就得到了足够多的收益差异数据，使用这些数据就可以对 β 进行估计。现在，我们要么计算 100 个投资组合的平均 β 值（时间平均），要么针对每个投资组合的 R_p 时间序列使用（A1）来对市场投资组合收益分别进行回归，由此可以得到 100 个投资组合的 β 值。

（ⅳ）在每一年中，个体股票的β值就取它所从属的“根据规模和β进行分组的”投资组合的β值。但这并不意味着个体公司股票的β值关于时间是常数。如果某家公司从“根据规模和β进行分组”的100个组合中某个组合转移到另外一个组合，那么为这家公司所指定的β值也会发生改变。

第二阶段就是针对2000家公司使用100个投资组合的β值来对（A2）进行横截面回归。对每个月（$t=1, 2, \cdots, T$）重复进行横截面回归，由此得到λ_0，λ_1和γ的时间序列，可以像正文一样来对它们进行分析。

$$\bar{R}_i = \lambda_0 + \lambda_1 \hat{\beta}_{pi} + \gamma Z_i + v_i \tag{A2}$$

这里的Z_i是横截面公司变量。

第9章

线性因子模型的应用

目的

* 说明如何使用市场（因子）模型来量化“事件”对股票收益的影响。

* 说明如何使用线性因子模型对共同基金、养老基金和套利基金的绩效进行评估。

* 考察运气和技巧两种因素在共同基金绩效评估中的重要性。

在这一章里，我们用前面叙述的因子模型来对实践中比较关心的两个问题进行考察。第一，如何度量“消息”或“事件”对股票收益的影响，这种影响可以看作是众多市场参与者对“消息”进行消化吸收后的反应。这里的消息包括：实际收益高于预期收益的声明或公司并购、合并等事件的宣告。如果市场有效，我们也拥有“合理”的股票收益模型，那么股票的异常收益可以反映这些“消息”对公司未来的影响。在有效市场上，今天的消息对未来的影响会迅速地被股票价格所吸收。第二，本章对过去的一些研究进行回顾，这些研究考察的是某个共同（和养老）基金在风险调整后的绩效是否更好一些，更好的绩效是否具有持久性，这些研究有助于投资者有效地选择共同基金之“星”，避免选择那些绩效特别差的基金。就共同基金的绩效评估而言，我们也使用自助法（bootstrap）来考察运气和技巧的作用。

9.1 事件研究

事件研究可以用来量化某一事件（或一组事件）对一家公司（或几家公司）价值的影响。这里的事件包括：并购事件的宣告、季报的发布、新股发行或首次公开发行（IPO）的公告。在美国，事件研究可以用来对内幕交易进行司法取证，取证过程需要对股票价格的异常波动进行度量。如果市场是有效的，那么事件对公司未来现金流的影响会迅速作用于股票价格，在一段时期之后就不会出现异常收益。

例如，对兼并与收购（M&A）的事件研究倾向于寻找如下目标公司，公司股票的异常收益平均为20%，平均来看，目标公司的股票价格对发起并购公司股票价格的影响接近于0。这意味着，平均来看，M&A不会降低竞争（否则，我们会看到目标公司和发起公司股价都会上涨），但发起者却要为获得“公司控制权”支付一大笔竞价费用。当然，对M&A的“事件研究”方法也存在其他形式，比如考察被并购公司在一个较长时期上（比方说5年）的绩效（如资本收益），再将其与行业中其他未被并购的类似公司进行比较。

事件研究方法也可用于检验公司发行新股或债券的宣告对公司价值的影响。这方面的研究表明：总体来看，发行新股会造成（异常）股价下跌2.5%至3.5%，而发行债券会导致大约0.2%的股价下跌，也就是说发行股票比发行债券的代价更大，这与Myers和Majluf（1984）

中资本结构的“啄序”（pecking order）理论是一致的（见 Cuthbertson 和 Nitzsche（2001a））。

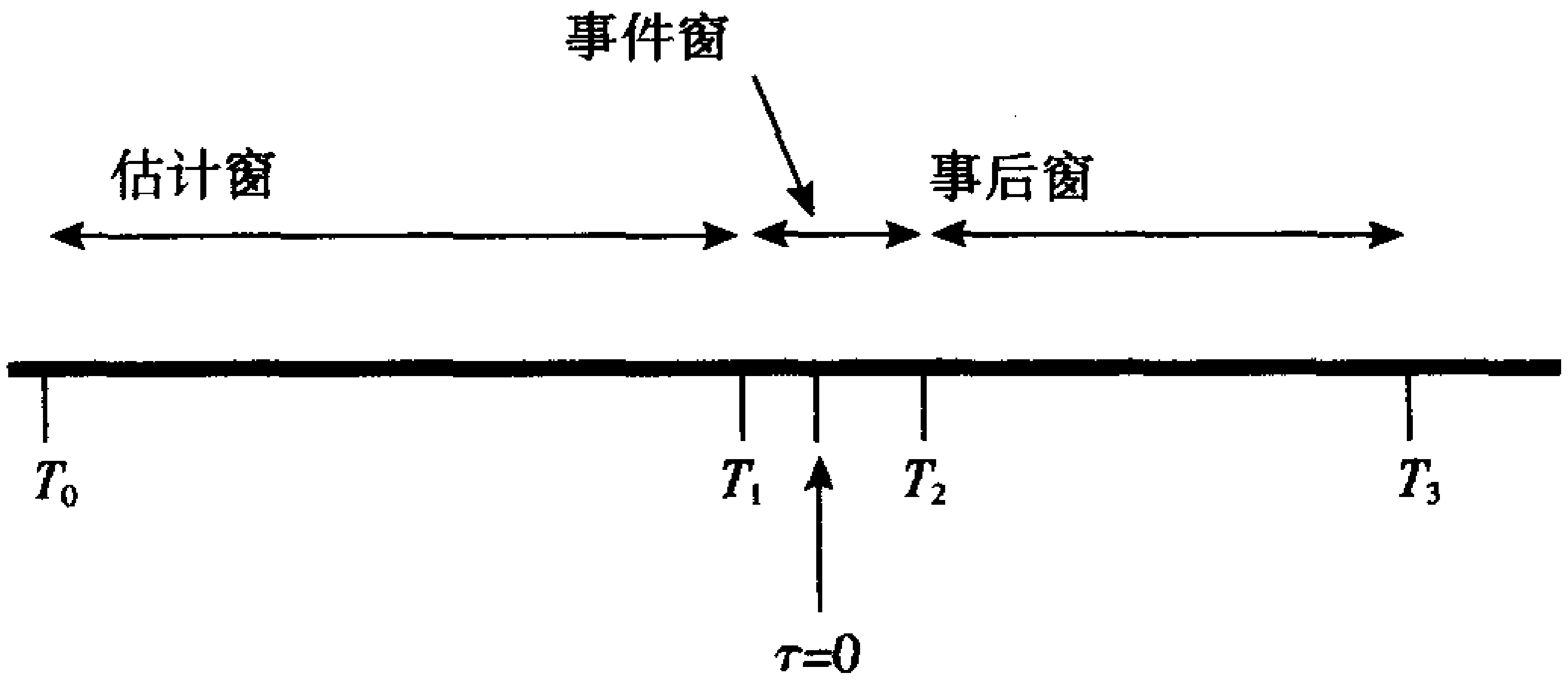


图 1 事件研究

实证研究中有两个与 CAPM 很“接近”的模型：单因子模型（SIM）和市场模型，实证研究经常用它们来“表示”股票收益。

$$R_{it}=\alpha_i+\beta_i R_{mt}+\varepsilon_{it}\quad t=T_0,\ \cdots,\ T_1$$

$$R_i=X_i\theta_i+\epsilon_i$$

其中，

$R_i=[R_{i,T_0+1},\ \cdots,\ R_{i,T_1}]'$ 为 $(T_1-T_0)\times 1$ 向量，

$X_i=[e1,\ R_m]$ 为 $(T_1-T_0)\times 2$ 向量，

$e1$ 是数字全为 1 的向量，

R_m 是代表市场收益的 $(T_1-T_0)\times 1$ 向量。

我们假设 $\varepsilon_{it}\sim niid\ (0,\ \sigma_i^2)$。市场模型允许 ε_i 和 ε_j 同时期相关（跨公司），而单因子模型则要求 $E(\varepsilon_i,\ \varepsilon_j)=0$。单因子模型在投资组合理论和风险度量（“在险价值”，Value at Risk）中非常有用，因为模型仅使用 β_i 和 σ_m^2 的估计值就可以计算出众多资产收益的方差和协方差（参见第 28 章）。市场模型与 CAPM 模型唯一的区别在于，市场模型将无风险利率排除在外，因为无风险利率的变动远远低于股票收益的波动，所以市场模型非常接近于 CAPM 模型。我们在这里用市场模型来阐述“事件研究”。

根据原理，事件研究方法是比较直观的。假定我们在时刻 τ 有一事件（图 1），如某一并购案的宣布。估计窗（T_0 至 T_1）用于度量目标公司股票的“正常收益”，通常使用的统计模型是市场模型。最小二乘法（OLS）得到的是一致、有效的估计值，即使 ε_i 和 ε_j 同时期相关。在实际运用中，当我们使用每日数据时，施加 $\beta_i=0$（即平均收益为常数的模型）的约束对结果几乎没有影响，因为在使用每日收益时，R_m 几乎不会增加方程的解释力。另外，我们也可以使用多因子统计模型（即右边的变量要更多一些），比如说 Fama 和 French（1995）的三因子模型、CAPM 模型（即使用超额收益且 $\alpha_i=0$）或者 APT，但它们在实际运用中的差别并不是很大。异常收益（AR，abnormal return）和事件窗 $[T_1+1,\ T_2]$ 上的累计异常收益（CAR，cumulative abnormal return）分别为

$$AR_{it}=R_{it}-\hat{R}_{it}=R_{it}-(\hat{\alpha}_i+\hat{\beta}_i R_{mt})\qquad (t\in[T_1+1,\ T_2])$$

$$AR_i=R_i^*-X_i^*\hat{\theta}_i$$

$$CAR_i=\sum_{t=T_1+1}^{T_2}AR_{it}=e1^{*\prime}AR_i$$

其中，

$R_i^*=[R_{i,T_1+1},\ \cdots,\ R_{i,T_2}]'$

$X_i^*=[e1,\ R_m^*]$

$R_m^*=[R_{m,T_1+1},\ \cdots,\ R_{m,T_2}]'$

$e1^*$是数字全为 1 的（$T_2 - T_1$）×1 向量

$\hat{\theta}_i = [\hat{\alpha}_i,\ \hat{\beta}_i]$

在无异常收益的零假设下（假设ε_{it}为 *niid*），

$$CAR_i \sim N(0,\ \sigma^2_{CAR_i})$$

其中，

$$\sigma^2_{CAR_i} = e1^{*\prime} V_i e1^*$$

$$V_i = [1 - X_i^* (X'_i X_i)^{-1} X_i^{*\prime}]\ \sigma_i^2$$

对 CAR 进行标准化后的 SCAR 为

$$SCAR_i = CAR_i / \hat{\sigma}^2_{CAR_i}$$

$SCAR_i$ 服从自由度为（$T_1 - T_0 - 2$）的学生 t 分布，可以用来对 CAR 为 0 的零假设进行检验。如果 N 家公司的事件窗都不存在重叠（比如说都在 $\tau = 0$ 发布公告），那么很容易对上述检验进行扩展，简单地将 N 家公司的 AR_i 和方差进行加总就可以了（参见 Campbell、Lo 和 MacKinlay（1997））

$$CAR_N = N^{-1} \sum_{i=1}^{N} CAR_i$$

$$\mathrm{Var}(CAR_N) = N^{-2} \sum_{i=1}^{N} \sigma^2_{CAR_i}$$

值得注意的是，因为有些事件研究涉及多家公司，所以在计算某个事件的 AR_i 和 CAR_i 时应将所有的样本公司进行分类。例如，我们在前面只考虑了目标公司的收益，显然也可以对发起公司的收益进行类似分析。类似地，当我们在对 N 家样本公司发布盈利公告这一事件的影响进行度量时，可以将样本公司分为三组：“好消息”、“坏消息”和“无消息”。“好消息”公司、“坏消息”公司和“无消息”公司分别是盈利“大于2.5%”、“小于 -2.5%”和“介于 ±2.5%之间”那些公司。这里的“消息”本身是用宣告盈利减去预期盈利来进行度量的，通过实际调查或对过去收益的时间序列进行简单外推就可以获得预期收益的数据。

大量证据表明：从事后来看，对可观测的公开事件而言，受到影响的公司股票价格沿着相同方向偏离初始位置的幅度大致相同。这些事件包括：盈利宣告（Bernard 和 Thomas（1989，1990）），红利派发（Michaely、Thaler 和 Womack（1995））、股票发行与回购（Loughran 和 Ritter（1995），Ikenberry、Lakonishok 和 Vermaelen（1995））和股票分析师推荐（Womack（1996））。事实上，根据 Chan、Jegadeesh 和 Lakonishok（1996）的观点，根据公司的历史盈利情况和与收益有关的“意外”事件或消息，股票收益在短期内（6 个月和 1 年）是可以预测的。

例如，Bernard 和 Thomas（1989）将纽约证券交易所（NYSE）和美国证券交易所（AMEX）的所有股票按照无法预料盈利（用盈利率的变化来进行度量，假定模型是随机游走模型）的大小进行分类，并对事件后 60 天进行了考察。他们发现，“盈利宣告的事后漂移效应”或“附带作用”确实存在，在所考察的 50 个季度中，“预料不到收益为正”的那一组（将所有公司分为 10 组）在 46 个季度中的表现比“预料不到收益较低的组”要好，两组的平均收益差异到达 4%——CAPM 模型的 β 值无法解释这一差异。更长时期的“附带效应”还包括红利派发（即一年期的异常收益）、股票的回购、首次发行或第二次发行（4～5 年的异常收益）等的效应。对更长时期的异常收益进行考察时存在一些统计问题，如果两家或多家公司的事件窗都长达 4～5 年，那么它们的事后收益就不再是独立的，这时就很难用标准方法进行解释（例如，参见 Barber 和 Lyon（1997）、Loughran 和 Ritter（2000）、Brav（2000））。

这种方法也可以用于对股票收益在短期内的动量特征进行考察。例如，Lasfer、Melnik 和 Thomas（2003）就用事件研究方法对 39 个国家（1989—1998）的股票市场每日价格总指数（如纽约证券交易所（NYSE）的全股指数）进行过研究，结果发现，如果股票指数在某一天出现大幅度变化，那么其后五天（即 $t = 1$ 到 5）的 CAR 都显著为正。$t = 0$ 时的事件定义为每

日收益变动超过之前时期（$t=-60$ 到 $t=-11$）平均波动率的两倍标准差，这一平均波动率在发达国家大约为 ±2.35%，新兴市场大约为 ±3.5%，“大的”价格变化在 3% 的时间内会发生。对正的大价格冲击而言，发达国家和新兴市场的指数在冲击后五日的 CAR 分别为 0.66% 和 1.54%；对于负的大价格冲击而言，CAR 则分别为 -0.14% 和 -0.50%。因此，在大的价格冲击之后，市场在短期内存在“动量特征”和“附带效应”。他们还发现，在流动性更差的市场上（用总市值来衡量）CAR 要更大一些；从时间来看，上世纪 90 年代后期的 CAR 要更小一些。然而，扣除交易成本后，可预测性和异常收益（相对于“世界”指数或每日预期收益而言）的存在并不必然意味着获利机会的存在。在有期货交易的发达市场中，交易成本和跟踪误差很低（小）是合理的，但在新兴市场上却不一定有这一结果，新兴市场上的期货交易成本相对较高，必要的指数数据也可能无法获得。在后面的章节中，我们将对“事件后漂移效应”这一普遍现象的可能原因进行考察（如代理人的学习、信息的缓慢扩散等等）。

一些实践问题

事件日可以是单独的一天（$\tau=0$），但是，如果事件（如并购公告）被提前泄露，那么事件窗就要向 T_1 的方向“往前移”，这样可以避免事件泄露对“正常收益”估计的不利影响。例如，在对 M&A 所进行的事件研究中，T_1 可以改为实际官方发布日的前 10 天（比如说）。

还有一个较为次要的问题是，计算每日收益时使用的是最后一笔交易的收盘价格，但是，对所有公司来说，事件不会在一天之类的相同时刻发生，对同一家公司而言，事件在不同日期中的发生时刻也不一样——这会导致“非交易”（non - trading）偏差。我们在前面所发展的统计结果都假定每日收益服从联合 *niid*。事实上这只是一种近似，因此我们要么依赖于渐进结果，要么对不同的参数分布进行蒙特卡罗模拟，要么使用自助法所得到的分布。

我们也可以使用非参数检验，尤其是符号检验（sign test）和秩检验，这些检验方法都不需要对分布进行假设。为简便起见，我们只涉及符号检验。如果总样本数为 N，那么将异常收益为正的个数记为 N^+。当 $N\to\infty$ 时，我们可以使用如下渐进有效统计量：

$$Z=\left[\frac{N^+}{N}-0.5\right]\frac{N^{1/2}}{0.5}\sim N\ (0,\ 1)$$

符号检验假设不同证券之间的异常收益独立，异常收益在零假设下的预期比例为 0.5。在事件研究中，这种检验方法可以作为 CAR 方法的一个补充。

9.2　共同基金的绩效

这一节首先讨论如何使用多因子模型来度量共同基金的风险暴露，如何用 Jensen α 来反映共同基金的异常收益，然后讨论近期英国和美国与共同（和养老）基金绩效有关的实证研究，包括对绩效持续性、表现最好的基金和表现最差的基金所进行的研究。最后关注一项试图考察如下问题的最新研究，绩效好的基金与绩效不尽如人意的基金之间的差异到底是因为运气不同还是因为基金经理的能力差异，也就是说，我们将对绩效经验分布的尾部进行考察。

无条件绩效度量指标

显然，在考察共同基金的绩效时，我们需要一些合适的方法来对风险进行调整。这一节介绍文献中最为常见的一些度量指标。Jensen（1968）用 α 来度量异常绩效，这种方法以市场模型——CAPM 模型——的设定为基础，可以将其表示为

$$R_{it}-r_t=\alpha_i+\beta_i\ (R_{mt}-r_t)\ +\varepsilon_{it} \tag{1}$$

其中，R_{it}是基金 i 在时期 t 中的收益，R_{mt}是市场投资组合的收益，r_t 表示无风险利率，实证研究中经常用短期国债利率代替无风险利率。如果 CAPM 模型是正确的均衡收益模型，那么投资

组合应位于证券市场线（SML）上，α 值也应该等于0。因此，在统计上显著的正值意味着基金经理具有超群的风险调整能力或选股技巧。对上面的方程（1）使用最小二乘法（OLS）就可以得到 α 的估计值——但我们经常用 GMM 来对标准误差进行修正。

Carhart 度量指标

Carhart（1997）的度量指标是四因子模型中的 α 估计值，四因子模型是在方程（1）中加入另外三个风险因子：规模、账面市值比和动量。可以将四因子模型表示为：

$$(R_{it}-r_t) = \alpha_i + \beta_{1i}(R_{mt}-r_t) + \beta_{2i}\mathrm{SMB}_t + \beta_{3i}\mathrm{HML}_t + \beta_{4i}\mathrm{PR1YR}_t + \varepsilon_{it} \tag{2}$$

其中，SMB_t、HML_t 和 $\mathrm{PR1YR}_t$ 分别表示规模、账面市值比和一年期动量效应三个风险因子的模拟组合（mimicking portfolios）。Carhart 的 α 度量指标就是方程（2）中的截距项。

四因子模型主要是建立在 Fama 和 French（1992，1993）及 Carhart（1995）对美国数据的经验研究结果基础之上。Fama 和 French 发现，在解释基金横截面平均收益时，包括市场收益、规模以及账面市值比三个风险因子的三因子模型，比仅使用超额市场收益的模型更有解释力。Fama 和 French（1992）还发现了股票收益与规模之间显著的负相关关系：投资规模较小的公司倾向于获得更高的平均收益。他们的研究根据规模来对股票进行排序，股票收益之间存在（平均来看）每月0.74%的收益差异。规模因子 SMB（“small minus big”）是小公司与大公司之间股票收益差异的一个度量指标。支撑规模风险因子设定的经济基础与相对预期有关。小公司的盈利预期对经济环境的变化更加敏感，在经济衰退时小公司遭受绩效下滑的可能性更大一些；还有一点就是，与大公司相比，小公司所体现的信息不对称要更为严重一些。这些因素都意味着规模具有风险载荷效应，投资者自然会要求一个更高的收益。

Fama 和 French（1992）也报告了股票横截面收益与账面市值比之间显著的正相关关系：高账面市值比公司股票的平均收益比低账面市值比公司高（作者所报告的结果是，两个收益之差达到每月1.5%）。账面市值比因子 HML（“high minus low”）是高账面市值比公司和低账面市值比公司之间收益差距的一个度量指标。横截面上的账面市值比很有可能是对公司前景预期的过度反应所造成的，高（低）账面市值比意味着公司的股票价格向下（向上）的调整幅度较大，从而账面市值比能预测横截面上的股票收益。

“动量”风险因子 PR1YR 捕获了 Jegadeesh 和 Titman（1993）的一年期动量异像效应，变量 PR1YR 是前期有着良好表现的股票（比如说涨幅榜最前面的1/5股票）和表现糟糕的股票（涨幅榜最后面的1/5股票）之间收益率差异的一个度量指标。对基金而言存在一种零投资策略，即买入前期有着良好表现的“动量”股票，卖空历史收益较低的股票，PR1YR 指标捕获了基金对零投资策略的敏感性。

条件绩效度量指标

前面所描述的 Jensen 和 Carhart 绩效度量指标都是无条件绩效度量指标。然而，无条件绩效度量指标确实与如下情景不相容：基金经理能识别预期收益和风险方面的市场信息变化并作出相应反映，重新构造投资组合。

Ferson 和 Schadt（1996）将 CAPM 模型扩展为条件绩效度量模型，在扩展后的模型中，t 时刻市场风险因子的因子载荷与由经济信息集 Z_t 所构成的工具变量向量线性相关：

$$\beta_i = \beta_{0i} + B_i'[z_{t-1}] \tag{3}$$

其中，z_{t-1} 是由 Z_{t-1} 对无条件均值的偏差所构成的向量，因此，β_{0i} 是条件 β 的无条件均值。将（3）代入（1）并对符号进行扩展，用 R_{bt} 表示基准投资组合（benchmark portfolio，在这种情形中就是市场投资组合）的收益，条件 β 类型的 CAPM 模型可以表示为：

$$(R_{it}-r_t) = \alpha_i + \beta_{0i}(R_{bt}-r_t) + B_i'[z_{t-1}(R_{bt}-r_t)] + \varepsilon_{it} \tag{4}$$

因为

$$E[z_{t-1}(R_{bt}-r_t)] = E[z_{t-1}]\,E[R_{bt}-r_t] + \mathrm{Cov}[z_{t-1},(R_{bt}-r_t)]$$

所以方程（4）捕获了择时交易变量z_{t-1}与基准投资组合预期收益之间的协方差。与前面一样，在异常绩效为零的零假设下$\alpha_i=0$。

方程（4）中的模型假设异常绩效度量指标为常数，然而，可能出现的情形是，异常收益也具有时变性（time varying）。Christopherson，Ferson和Glassman（1998）假设条件α是工具变量z_{t-1}的线性函数：

$$\alpha_i=\alpha_{0i}+A_i'\ [z_{t-1}] \tag{5}$$

将（5）代入（4）就得到

$$(R_{it}-r_t)\ =\alpha_{0i}+A_i'\ [z_{t-1}]\ +\beta_{0i}\ (R_{bt}-r_t)\ +B_i'\ [z_{t-1}\ (R_{bt}-r_t)]\ +\varepsilon_{it} \tag{6}$$

条件α方法也可以用于Carhart四因子模型。Ferson和Schadt（1996）及Christopherson、Ferson和Glassman（1998）将经济信息集合Z_t作为工具变量，以前的研究表明这些工具变量对时变证券收益和风险具有一定的预测功能，工具变量包括：一月期国库券收益的滞后值、市场因子红利收益的滞后值（the lagged dividend yield of the market factor）、期限结构中斜率的滞后值、不同级别公司债券收益率差的滞后值和能捕获一月效应的虚拟变量。

A. 英国共同基金

Quigley和Sinquefield（1999）对英国1978—1997年间所有单位信托基金（即开放基金）——总共752只，包括一些已经不存在的基金——进行了考察。他们根据竞买竞卖价格（bid－price－to－bid－price）来计算收益（即基金投资者所支付的价格已经包含了买卖价差），所有红利收益再投资于基金。这反映了作者的目的是为了对基金经理的绩效进行评估，而不是对基金购买者的收益进行考察。

构造一个由所有单位信托基金（存活的和已经退市的都包括在内）组成的等权重投资组合，然后对Fama－French类型的三因子模型进行估计，结果发现，在5%的显著性水平下，α值为每月－0.09个基点。四种不同投资类型的基金（即投资于成长型股票、收益型股票、普通股票和小公司股票等四种类型的基金）都出现如此糟糕的绩效。对每种类型的基金而言，由这种类型所有单位信托基金构成的等权重投资组合的α值也为负，其中绩效最差的是那些投资于小公司股票和收益型股票的基金。

绩效持续性

α的历史值反映了基金过去的平均异常绩效，我们感兴趣的是这些绩效是否可以持续，也就是说，如果我们在时间t按照某种标准（比如说选择过去一年表现最好的1/10股票）来构造一个股票投资组合，持有这个组合直到时间$t+k$再进行调整，那么在对风险进行调整的基础上，这种“递归构造（recursively formed）”投资组合是否可以取得良好绩效?

Quigley和Sinquefield（1999）最先对基金绩效持续性进行了考察，对每一年而言，根据过去12个月的收益对基金进行排序，将所有基金分为10个组，每组内的基金按照等权重方法来构造投资组合。这10个投资组合被称为绩效投资组合，每个投资组合持有一年后再进行调整。每年都重复这种策略，由此可以得到10个投资组合收益的时间序列。结果表明，绩效最好与最差基金之间的（复利）年收益相差3.54%（即零投资组合（a zero－investment portifoio）的收益）。运用这种策略似乎很容易“战胜市场”，但在实际运用这种策略时，绩效最好投资组合每年都要对80%的基金进行调整，竞买竞卖所付出的成本高达5%，异常收益很有可能被抵消掉。此外，如果用三因子模型来对这些绩效投资组合的风险进行调整，那么10个投资组合的α值表明绩效持续性不显著。两个最好的“绩效投资组合”的α值尽管为正，但在5%的显著性水平下还是不显著。与此相对照的是，四个最差绩效投资组合的α值为负，且在统计上都显著。这与针对美国市场的实证研究结果（即Carhart（1997）和Malkiel（1995））相吻合，也就是说，追求持续性的策略涉及较高的换手率，在对风险进行调整后，绩效较差的投资组合策略具有持续性，而绩效较好的投资组合策略则没有持续性。

Quigley 和 Sinquefield（1999）还根据过去三年的数据（递归使用）按照三因子模型中的值对基金进行排序并将所有基金分为 10 组，所得到的定量结果与上述结果类似：在 10 个“绩效投资组合”的 α 值中，只有最后两个绩效投资组合的 α 值在统计上显著且符号为负。这再一次说明了如下结论：基金绩效的良好表现总是无法持续，而基金绩效的拙劣表现却具有持续性。使用同样方法将持有期扩展到一年以上，他们发现任何形式的绩效持续性在三年后都消失了，根据 α 进行排序的调整前后顺序之间的相关性降为 0. 12。

Blake 和 Timmermann（1998）是对 9 个英国共同基金绩效进行研究的另一篇重要文献，他们对 2 375 只共同基金进行了考察，有些基金并不是只投资于英国证券市场，已经退市的基金有 973 只，样本期选择为 1972 年 2 月至 1995 年 6 月，略早于 Quigley 和 Sinquefield（1999）中的样本期。

然而，在绩效持续性研究方面，Blake 和 Timmermann（1998）的结论与 Quigley 和 Sinquefield（1999）在很多方面都不同。无条件因子模型使用三个解释变量：股票市场指数的超额收益（$R_{mt}-r_t$）、小市值公司股票收益 R_{st} 相对于市场指数收益 R_{mt} 的超额收益（$R_{st}-R_{mt}$）、小市值公司股票收益 R_{st} 相对于英国五年期国债收益 r_t 的超额收益（$R_{st}-r_t$）：

$$(R_{it}-r_t)=\alpha_i+\beta_{mi}(R_{mt}-r_t)+\beta_{sm,i}(R_{st}-R_{mt})+\beta_{sr,i}(R_{st}-r_t)+\varepsilon_{it} \tag{7}$$

Blake 和 Timmermann 在对英国基金绩效持续性进行研究时将分析限制在股票型基金和平衡型基金上，也就是说，对这些基金而言，方程（7）中右边变量都是很好的基准性变量。这种基金共有 855 只。根据前 24 个月的“历史 α 值”来对基金进行排序，将基金分为四组，取前后两组分别构造两个等权重投资组合，两个投资组合的持有时间都只有 1 个月，每个月对组合进行一次调整。文章对五种投资风格的基金执行了上述程序：投资于成长型股票、收益型股票、普通股票、小公司股票的基金和平衡型基金。除平衡型基金外，由每种风格前 1/4 基金所构造的递归投资组合在样本期内的异常收益为正。由每种风格后 1/4 基金所构造的递归投资组合在样本期内的异常收益为负。（这一结果对投资组合权重的选择而言是稳健的。）结果意味着，绩效最好和绩效最差的基金具有绩效持续性。研究还表明，投资于成长型股票和小公司股票的基金的绩效持续性在统计上显著。

在这篇文章中，绩效最好基金的绩效持续性与 Quigley 和 Sinquefield（1999）中的结果截然相反，而对绩效最差的基金而言，两篇文章中的绩效持续性结果却一致。对两篇的分析进行比较是很困难的：首先，他们对异常收益或风险调整后收益的度量方法不同；其次，他们对投资组合进行调整的频率不同。Quigley 和 Sinquefield（1999）在每一年中根据过去三年的异常收益来对投资组合进行调整，而 Blake 和 Timmermann（1998）则根据过去两年的异常收益每月都对投资组合进行调整。Blake 和 Timmermann（1998）每月对投资组合进行调整的策略比每年才进行一次调整的策略更加理性，每月调整考虑了基金在一年中可能退市的情形。尽管存在这种优势，那些想利用这种持续性“异像”的投资者在每一次调整时都要承担大量的重构工作，承担数量可观的交易成本和管理费用。异常收益很有可能被抵消掉。因为在 Blake 和 Timmermann（1998）的分析中，收益是在竞买竞卖价格基础上计算出来的，所以收益是包含交易成本的总收益。另外，他们的论文没有提及在逐月构造绩效投资组合过程中转换成本方面的信息。因此，试图利用持续性异像这一策略的投资者对成本的推断将是非常困难的。

在 Quigley 和 Sinquefield（1999）、Blake 和 Timmermann（1998）及 Fletcher（1997）所描述的绩效持续性检验中（通过递归投资组合的构造来对绩效持续性进行检验），一个困难是要对大量数据进行加总，而不是针对个体基金数据来考察绩效持续性问题。Lunde、Timmermann 和 Blake（1999）提到了这一问题，Allen 和 Tan（1999）也对这一问题进行过考察，他们对英国投资信托基金（即封闭式基金）而不是单位信托基金的绩效进行了评估。Lunde、Timmermann 和 Blake（1999）首先将英国股票型共同基金的收益分为四组，收益的调整指标与 Blake 和 Timmermann（1998）类似。基金分组的标准是过去 36 个月的收益，根据随后 36 个月的表现，记录下每组基金落入某个给定组的比例，由此得到各组之间的转移概率。绩效最差的基金

（最后一组）在36个月后仍然在这个组中的概率为0.332，绩效最好的基金（最前面的一组）仍然在这个组中的概率为0.355。如果只考虑存活的基金，那么这两个概率值分别降为0.284和0.317。在不存在持续性的零假设下，所有转移概率都应等于0.25。因此，Lunde等对包括存活的和已退市的所有基金进行研究的结果明显可以拒绝这个零假设。

上述研究（见表1）涵盖了英国共同基金绩效、异常收益和绩效持续性文献中所出现的主要问题和结果。还有一小部分文献使用前面描述的程序对英国共同基金业进行研究，报告的结果与上述结果类似。下面对这些文献进行简要回顾。

WM公司（1999）对1979—1998年期间英国存活的收益型和成长型基金进行了考察，使用未加工的收益数据来对绩效最好的1/4基金在五年中的绩效持续性进行了考察。并不奇怪的是，绩效最好的那一组中的基金在下一个时期仍然在这一组中的比率急剧下降。如果样本期不重叠，那么上述比率比根据概率所预测的值小。这一结果说明，历史上的相对绩效不是未来绩效的一个好的预测指标。

除了对英国单位信托基金和投资信托基金的绩效进行考察之外，也有研究对其他类型基金，尤其是养老基金的绩效进行了考察。Blake、Lehman和Timmermann（1999）对1986—1994年期间的306只养老基金进行了研究，研究特色之处在于：相对于单位信托基金而言，对养老基金的研究很难得出相对绩效和绩效持续性的相关结论。养老基金市场更加关注投资政策，基金的绩效由投资组合中资产的组合方式来决定，而不是由基金经理的选股技巧和择时交易策略来决定。也就是说，相对于其他类型的基金而言，养老基金的基金经理并不采用主动型交易策略，Blake等所报告的一个令人感到惊讶的结果是，资产配置策略、选股技巧和择时交易策略等在横截面上的收益差异非常小。Brown、Draper和McKenzie（1997）对600只养老基金进行过研究，得出的结论与之相似。

表1　英国共同基金绩效研究

文献	样本期	基金类型	是否对收益进行风险调整	是否对交易成本进行控制	是否对基金存活状况进行控制	绩效持续性
Allen和Tan（1999）	1989～1995	131只投资信托基金	是	否	进行间接检验	有，最好和最差的基金都存在
Blake和Timmermann（1998）	1972～1995	2375只单位信托基金，分为20种风格（在绩效持续性检验中使用855只股票型基金）	是	否	是	有
Fletcher（1997）	1980～1989	101只单位信托基金：成长型、收益型和普通型	是	否	基金至少有两年的收益值	无
Leger（1997）	1974～1993	72只投资信托基金	是	否	否	无
Lunde、Timmermann和Blake（1999）	1972～1995	与Blake和Timmermann（1998）相同	是	否	是	有
Quigley和Sinquefield（1999）	1978～1997	752只单位信托基金：成长型、收益型、普通型和小公司型	是	一部分	是	绩效较差的基金存在，绩效较好的基金不存在
WM公司（1999）	1979～1998	英国收益型与成长型基金	否	否	否	无

表 2　　　　　　　　　　美国共同基金绩效研究

文献	样本期	基金类型	是否对收益进行风险调整	是否对交易成本进行控制	是否对基金存活状况进行控制	绩效持续性
Carhart（1997）	1963～1993	所有股票型共同基金	是	是	是	绩效好的基金没有，绩效差的基金有
Chen、Jegadeesh 和 Wermers（2000）	1975～1995	所有共同基金（主要是股票型基金）	是	使用总收益率	是	有，但主要原因是动量效应，而不是选股技巧[1]
Christopherson、Ferson 和 Glassman（1998）	1979～1990	273 只股票型养老基金	是	扣除交易佣金，不扣除管理费用	否	有，集中于绩效较差的基金
Elton、Gruber 和 Blake（1996）	1977～1993	188 只股票型共同基金	是	是	是	有
Fletcher（1999）	1985～1996	85 只投资于美国市场的英国单位信托基金	是	扣除管理费用，但包括佣金	是	无
Goetzmann 和 Ibbotson（1994）	1976～1988	728 只共同基金	是	否	无	有
Grinblatt 和 Titman（1992）	1975～1984	279 只共同基金	是	是	无	有，在绩效较差的基金中证据最为明显
Hendricks、Patel 和 Zeckhauser（1993）	1975～1988	无佣金，以增长为导向的股票型共同基金	排序时为否，估计时为是	扣除管理费用	是	短期有，更多地表现在绩效较差的基金上
Malkiel（1995）	1971～1991	所有股票型共同基金	是	是	是	70 年代有，80 年代无
Volkman 和 Wohar（1995）	1980～1989	332 只共同基金	是	是	是	有

注：该研究对每只基金的买卖交易策略进行了考察，结果发现，绩效较好基金被动持有股票（从一个时期到下一个时期）的表现比绩效较差基金被动持有股票的表现更好一些。然而，绩效较好基金新购买股票的表现只比绩效较差基金新购买股票的表现好一丁点。这表明是动量效应而不是选股技巧能解释绩效持续性。

为了得出英国共同基金绩效持续性的整体结论，将众多结果进行接合比较是很困难的，原因有多个方面。不同研究考察的是不同的样本期，样本长度也不一致。不同研究使用不同的风险调整或异常收益度量指标。此外，不同研究的分组标准不尽相同，有的分为十组，有的分为五组，有的分为四组，有的根据中值来进行分组。对绩效持续性的考察时期和排序方法也不一样，如所选择的收益期可能为一年、两年、三年或五年。有些研究使用加总数据来考察绩效持续性（参见 Quigley 和 Sinquefield（1999）、Blake 和 Timmermann（1998）），有些研究则使用个体基金层面的数据。最后，有些研究对基金存活变量进行了控制，致力于扣除交易成本后的绩效异像研究，有些研究则没有考虑这些因素。

尽管现有研究存在上述种种差异，英国共同基金业的绩效还是存在如下特征：对绩效最好的基金而言，绩效持续性效应要弱一些。在考虑基金投资者或客户的交易成本后运用基金绩效的异像特征来进行获利的投资策略是值得怀疑的。与此相对应的是，绩效较差基金的持续性证据要更为确凿一些。因此，从整体来看，基金绩效持续性分析对投资者如何避免那些绩效较差的基金还是具有一定的帮助，但对如何选择基金却没有多大的帮助。

B. 美国共同基金

与对英国和欧洲共同基金的研究相比，对美国共同基金进行研究的文献要多得多，最新研究所呈现的核心问题如下（见表2）。

Carhart（1997）认为，许多早期研究发现的结果在很大程度上是由基金中股票的“动量效应”所导致的，基金的绩效应归功于基金经理的好运气而不是选股技巧。Carhart（1997）是一篇重要的、全面的、广为引用的文献，我在这里对其进行详细讨论。Carhart 在 1963—1993 年的样本期上使用迭代投资组合构造方法，根据滞后一期未经调整的收益将所有基金进行分成十组，将每组内的基金使用等权重方法来构造投资组合，持有投资组合一年后再对其进行调整，Carhart 所使用的基金收益是剔除掉运营费用和交易成本的净收益。

Carhart 首先针对上述十个“绩效投资组合”使用 CAPM 模型来估计 Jensenα 的值。然而，CAPM 模型明显无法解释这些（十个）投资组合在横截面上的平均收益，绩效最好和最差绩效投资组合的 CAPM -β 值几乎相等，因此所得到的 Jensen 值与简单收益有着相同的分散程度。Jensenα 的估计值表明，在绩效最好的绩效投资组合中，每月存在22个基点的异常收益；在绩效最差的绩效投资组合中，每月存在 -45 个基点的异常收益；许多绩效组合的 α 值（在5%的置信水平下）在统计上显著。

与仅使用超额收益的情形相比（即 CAPM 模型），在 Carhart 的报告中，四因子模型可以解释十个“绩效投资组合”平均收益差异的很大一部分，而规模因子（SMB）和（PR1YR）动量因子能解释收益在横截面上的绝大部分变化，绩效最好和绩效最差两个绩效投资组合之间收益差的一半可以用动量因子来解释。然而，对所有绩效投资组合而言，四因子模型所估计的 α 值为负，绩效最差的三个绩效投资组合所对应的 α 值在统计上显著。这就导致了如下结论：对风险进行调整后的净收益只能为绩效差的基金的持续性提供证据，研究结果不支持基金经理投资技巧起作用的结论。

持续性

Carhart 对基金绩效持续性也进行了分析，根据四因子模型所估计的“历史”α 值（使用过去三年的数据进行递归估计）将基金分为十组，然后每年进行一次调整。四因子模型可以解释十个“绩效投资组合”的横截面收益，在绩效最好的绩效投资组合中，绩效持续性不显著，与此相对照的是，绩效最差的绩效投资组合中却存在绩效持续性。

Carhart（1997）还进一步考察了动量效应，正如前面描述的那样，动量因子能解释绩效最好和绩效最差两个绩效投资组合收益之差的一半。股票遵循动量策略的那些基金一直都可以获得比平均收益还要高的收益，而上述四因子模型所得到的结论是，在对风险进行调整后，具有这种表现的基金不存在异常收益。为了进一步考察运用动量策略的基金是否具有一致的良好表现，文章根据动量因子载荷（factor loading，PR1YR 载荷）来对所有基金进行排序，在排序后时期中，一年期动量效应“最显著”的基金却表现得不是很好（用 Carhartα 值来进行度量）。Carhart 分析中所使用的收益是将交易成本和相应费用扣除后的净收益，他对这个显而易见的矛盾进行了如下解释：拥有高动量因子载荷的基金不会积极地遵循动量策略，而是“随意地”[Carhart *p*.73] 持有上一年度表现较好的股票。通过简单地持有这些表现较好的股票，基金在不承担滚动动量策略的额外交易成本的情况下就可以获得一个更高的平均收益。

Carhart（1997）试图解释异常收益的决定因素，异常绩效用前三年数据所估计的单个基金 α 值来进行度量（用四因子模型中的 α 值来进行度量），以月为基础进行滚动回归：

$$\begin{aligned}\alpha_{it} = (\bar{R}_{it} - \bar{r}_t) &- b_{1i,t-1}(\bar{R}_m - \bar{r}_t) - b_{2i,t-1}(\overline{\mathrm{SMB}_t}) \\ &- b_{3i,t-1}(\overline{\mathrm{HML}_t}) - b_{4i,t-1}(\overline{\mathrm{PR1YR}_t})\end{aligned} \tag{8}$$

然后，针对每一个月进行横截面回归：

$$\alpha_{it} = \delta_1 + \delta_2 (X_{it}) + \varepsilon_{it} \tag{9}$$

X_{it}在这里表示基金的特征：当月的基金规模（对数值）、费用（基金管理费）和股票交易费用。与 Fama 和 Macbeth（1973）一样，Carhart（1997）先对方程（9）中的横截面关系进行了估计，然后再将“月度”估计系数进行平均。他所得到的结论是基金绩效与规模之外的所有特征变量之间具有很强的负相关关系。尤其有趣的是费用比率的系数：费率每提高 100 个基本点，年度异常收益（即横截面上的 α 值）下降 1.54%。因此，基金的这些特征似乎能进一步解释异常收益和绩效较差基金中的绩效持续性。

Chen、Jegadeesh 和 Wermers（2000）回到 Carhart（1997）所考察的由动量效应所导致的持续性质问题。Chen、Jegadeesh 和 Wermers（2000）的论文为如下问题提出了更为深刻的见解：“绩效好”的基金是真的遵循动量策略，还是如 Carhart 所说的只是简单地、随意地持有前期表现良好的股票。Chen、Jegadeesh 和 Wermers（2000）的数据集合包含 1975—1995 年期间“绩效好”和“绩效差”共同基金（根据未经调整的季度收益将基金分为 5 组，选取绩效最好的一组和绩效最差的一组）的股票买卖信息。

文章首先报告了如下结论：绩效好的基金在排序期前两个季度中的平均收益明显比绩效差的基金的平均收益高。文章一开始就指出了绩效持续性的相关结果。在对基金买卖交易的分析中，文章发现：“绩效”基金所“购买”股票的历史收益明显比“绩效差”基金所“购买”股票的收益高。这似乎是说，与绩效差的基金相比，绩效好的基金应遵循动量投资策略，这与 Carhart（1997）所提到的“绩效好”基金并不遵循动量策略而是随意地持有前期表现良好股票的观点相反。不过，文章也发现：绩效较差基金所购买股票的收益比历史收益高，未来收益比绩效较差基金所持有股票的整体收益高，这为 Carhart 的观点提供了支持，也就是说，如果绩效较差的基金继续持有过去表现不好的股票，那么基金未来的整体收益会降低。然而，在一年的持有期中，不管对收益是否进行风险调整，绩效好的基金所购买的股票的表现不会显著地比绩效差的基金所购买的股票的表现好。这些结果表明：绩效好的基金并不具有更好的选股技巧，但能从所持有的股票组合的动量效应中获得收益。然而，这一效应很快就会消失，与绝大多数研究的结果一样，这种持续性——当持续性存在时——不会持续太久。

基金特征与绩效

Wermers（2000）也对如下问题进行了考察：在主动型基金中，交易活动的水平或基金内股票的转换交易量与基金绩效之间是否存在联系。为了强调这个问题，Wermers（2000）使用 Carhart（1997）、Blake 和 Timmermann（1998）等所描述的递归投资组合构造方法，但他的分类标准有所不同，他根据基金在上一年中的转换交易量来对所有基金进行排序，将所有基金分为十组并构造十个基金组合（也可以将所有基金分为 5 组），投资组合的持有期为一年。这种方法可以用来检验转换交易量最大的基金组合是否与最好的基金绩效相一致。

Wermers 在对毛收益率（包含交易成本与运营费）进行考察后发现，在 1975—1993 年的样本期内，最上面基金组合的年收益比最下面基金组合高 4.3%，结果在 10% 的显著性水平下显著。Wermers 也对这一差异的原因进行了考察，按照原因的重要性进行递减排序，依次为基金的投资风格、选股能力和择时交易能力。如果考虑净收益率，那么十个投资组合中最上面与最下面两个组合的收益差异降至 2.1%（但不显著），如果将基金分成五组，那么最上面与最下面两个组合的收益差异为 2.7%（5% 的显著性水平下显著）。如果对净收益进行风险调整（Carhart 值），那么最上面与最下面两个组合的绩效之间没有差异。

并不奇怪的是，Wermers 报告了如下结果：基金的转换交易量与交易成本、基金费率之间存在显著而又持久的相关性。非常有趣的是，我们可以在 Wermers（2000）研究的基础上得出如下结论：基金的交易成本和运营费率与基金绩效之间存在类似关系，就像前面所描述的转换交易量与基金绩效之间的关系一样。总的说来，Carhart（1997）中绩效（至少是对风险进行调整后的净绩效）与转换交易量和运营费用之间的关系比 Wermers（2000）中的负相关关系要强得多。

Chalmers、Edelen 和 Kadlec（1999）对成本/绩效问题的考察更为全面深入，用多种绩效度量指标对1984—1991年期间基金交易成本、基金费率、转换交易量与基金绩效之间的关系进行了考察。在每一个季度内，根据总成本对所有基金进行排序，将所有基金分为五组，计算每组样本的季度平均绩效，分别以交易成本、费用比率、转换交易量为分类标准重复这个程序。在根据三种成本标准进行排序和分组的情形中，成本与绩效之间存在强（且显著的）负相关关系，转换交易量与绩效之间的关系也为负，但不显著。Chalmers、Edelen 和 Kadlec（1999）的结果与 Carhart（1997）大体一致。

对基金有无佣金（认购费）的相对绩效问题进行考察的文献有：Ippolito（1989）、Elton、Gruber、Das 和 Hlavka（1993）、Grinblatt 和 Titman（1994）、Droms 和 Walker（1994）、Gruber（1996）、Fortin 和 Michelson（1995）及 Morey（2003）。Ippolito（1989）的研究认为收取佣金基金为投资者赚取的收益抵消了发行时的佣金，其他研究则普遍认为基金有无佣金在绩效上没有显著差异。

Morey（2003）也许是对这一问题考察得最为全面的文献，因为他是唯一一个对基金绩效用佣金来进行调整的研究者，无论基金是否收取佣金，都要对基金的绩效进行调整。此外，Morey（2003）还对佣金高低与相对绩效之间的关系进行了考察，使用了多种相对绩效度量指标。有的指标对风险进行了调整，有的指标就是原始收益，也包括 Jensen 度量指标。Morey 发现，在用佣金对收益数据进行调整之后，无论使用哪种度量指标，无佣金基金的绩效都比收取佣金基金的绩效好。高佣金与低佣金基金之间的绩效差异几乎不显著。

投资流动与绩效

共同基金绩效研究的一个更为深入领域是对基金绩效与基金投资的资金流入流出之间的关系进行考察。资金流动与基金绩效之间的因果关系方向一直存在争论。许多研究对资金流动与绩效之间的正相关关系进行了检验，如 Gruber（1996）和 Zheng（1999），而 Edelen（1999）却对相反的因果关系进行了检验。

现金流动定义为基金总的净资产与期初资产之差的改变量，Gruber（1996）根据 $t-1$ 期的基金绩效将所有基金分为十组来对期中的资金流动进行考察，然后对 t 滚动使用这一程序。Gruber 所报告的结果是，基金绩效与资金流动之间是一种强而显著的关系。

类似地，将基金现金流动对滞后一年及两年的绩效指标进行回归的结果也说明基金绩效对后续资金流动具有显著影响。然而，更为有趣的是，这一程序也为投资者使用什么绩效度量指标对基金经理进行评价提供了见解。尽管许多绩效度量指标都证明了绩效是资金流入/流出的显著原因，也许令人感到惊讶的是，Gruber（1996）还是认为使用四因子模型中的 α 绩效度量指标得到的结果最为显著和稳健。这似乎说明，尽管面临着市场、规模、价值、信用风险等各种因素，投资者在选择投资对象时还是十分老练。这也说明基金评级公司在对基金进行排序时也遵循这样的准则，这就是投资者倾向于选择高 α 值基金的原因。DelGuercio 和 Tkac（2000）也发现基金 Jensen α 值与后续资金流动之间的显著关系，共同基金和养老基金都有这样的特征，尽管对养老基金而言，跟踪误差（tracking error）对资金流动的影响也很重要。（跟踪误差是分散风险的一种度量，通常用投资组合对基准指数偏差的波动率来进行度量。）

在发现良好（糟糕）绩效好导致了后续资本流入（流出）的证据后，Gruber（1996）继续考察如下问题：投资者对资本投向的重新定位是否可以改进投资绩效。将 t 期资金流入乘以后续各期（多个时期）用风险调整后的基金收益，对所有基金和所有时期进行加总，再将结果表示为所有基金总资本流入的百分比形式。结果发现，“新资金”经风险调整后的平均收益每年为29个基点。对资金流出使用类似程序就可以对投资者从基金中提取资金能节省多少钱进行度量，结果表明，投资者每年能节约22个基点。

一个较为宽泛的类似研究是 Zheng（1999），他对如下问题进行了考察：资金流动或“新的货币流动”对未来收益是否有预测作用，总的货币资金流动是否为投资者提供了获得异常

收益的充足信息，也就是说，基金的资金流动是否为精明投资策略提供了信号。Zheng（1999）根据基金的资金流动实施了一系列的交易策略，并这些交易策略是否能获得异常收益进行了检验。例如，一种策略是投资者持有资金流入为正的基金组合，每一季度都要对投资组合进行更新；另外一种交易策略是持有现金流为负的投资组合。也可以是如下情形：一种策略是持有高于平均现金流的基金组合，另一种策略是持有低于平均现金流的基金组合。在所有情形中，每一季度都要对基金组合进行调整，由此可以得到各个投资组合收益的时间序列。可以使用单因子和多因子模型中的 α 绩效度量指标来对这些“以货币流动为基础的基金”进行评估，货币流动或资金流动的定义参见 Gruber（1996）。Zheng（1999）对 1970—1993 年的样本期进行了考察，结果发现，有资金流入基金的绩效显著地优于那些有资金流出的基金，这一结果与 Gruber（1996）基本一致。然而，相对绩效不会持续太久。另外，从整体来看，有资金流入的基金并没有显著地“击败市场”。

对于 Gruber（1996）和 Zheng（1999）的研究，人们应注意的是，资金流动与绩效的这种关系也许仅仅是因为基金绩效的持续性（如果存在的话），也就是说，基金的优良绩效吸引了资本或“新资金”。因为基金绩效具有持续性，投资者在这些“新资金”上的投资反过来又能获得高额收益。Sirri 和 Tufano（1993）对单个基金的资金流动进行了深入研究，大致结果是，过去绩效最好的那些基金有资金流入，而绩效最差的那些基金也没有资金流出。这可能反映了如下情形：（i）基金业获得了巨大的发展，过去收益高的基金引了大量的新增资本，绝大部分资金流动是新增资本的流动而不是基金之间的资金流动；（ii）投资者将资金从一只基金转移到另一只基金要承担一定的成本。

Edelen（1999）对资金流动与绩效之间关系的研究得出的是截然不同的结论。Edelen 假设两者之间存在与 Gruber（1996）相反的因果关系，特别地，他认为资金流动会对后续基金绩效产生负面效应。根据共同基金所提供的两种服务，他将共同基金的交易策略分为两类。基金经理在一些信息——这些信息能导致正的“风险调整”后的收益——的基础上采取谨慎（理性）的交易策略。另外，基金经理的主要职责（和目的）是满足投资者的流动性需求，也就是说，以低成本为投资者提供流动性权益头寸。Edelen（1999）认为，资金流动或流动性冲击迫使基金经理采取“由流动性所驱动的交易策略”，而这种策略不是谨慎交易策略。经历资金流动冲击的基金会立即改变现金/权益资产的持有比例，改变基金的目标组合，足够大的资金流动规模将使基金的现金头寸发生很大的变化。首先，这使得投资者的风险收益权衡决策变得很复杂；其次，这使得基金经理与指数追踪目标或击败市场指数目标相妥协。因此，为投资者提供流动性权益头寸的做法使得共同基金经理进行一些边际权衡交易。Edelen 认为，共同基金经理出于流动性考虑所进行的一些交易在市场中扮演了噪声交易的角色，如果噪声交易者面临预期损失，那么基金负的收益绩效与资金流动规模成比例。

Edelen（1999）将共同基金的异常收益分为两部分：由流动性驱动交易所带来的收益和由信息驱动（谨慎）交易所带来的收益。文章首先对由流动性驱动的交易进行了估计：

$$f_{i,t} = c^I f_{i,t}^I + c^0 f_{i,t}^0$$

$f_{i,t}^I$和$f_{i,t}^0$在这里分别表示基金 i 在 t 时刻的资金流入量与流出量，将基金的股票购买量$f_{i,t}^I$对进行回归，将基金的股票卖出量对$f_{i,t}^0$进行回归，由此得到的系数分别为 c^I 和 c^0，即 c^I 和 c^0 分别表示“资金流动性交易”的反应系数。接下来，作者构建了如下回归方程：

$$AR_{i,t} = \lambda f_{i,t} + \delta d_{i,t} + \varepsilon_{it}$$

$AR_{i,t}$表示基金 i 在时期 t 中的异常收益，$f_{i,t}$表示所估计的由流动性驱动的交易，$d_{i,t}$表示所估计的由信息驱动的（谨慎）交易，它等于基金股票买卖总交易量减去$f_{i,t}$，$\varepsilon_{i,t}$是随机扰动项。为了将绩效持续性纳入到考察范围之内，作者在回归方程的右边增加了 $AR_{i,t}$的滞后值。最后，为了避免 Gruber（1996）中所出现的由 $AR_{i,t}$和$f_{i,t}$之间的逆向因果所导致的推断问题，作者以$f_{i,t}$的滞后值作为工具变量。尽管困难重重，Edelen（1999）的估计结果仍然在一定程度上为“资金流动对基金绩效有负面影响”的假设提供了证据。

将 Gruber（1996）、Zheng（1999）、Sirri 和 Tufano（1993）和 Edelen（1999）两个方面的研究结合起来，我们大致可以得出这样的结论：较好的基金绩效吸引了资本，但为投资者提供流动性投资需求的压力使得基金经理无法对权益资金进行最优配置，也就是说，基金经理无法对基于信息的理性交易作出迅速决定，这导致了基金后续绩效的降低。

对共同基金资金流动进行深入研究的一个渠道是对总量资金流动是否对市场收益产生影响的问题进行考察。Warther（1995）对这一问题进行了考察，在此简要地叙述相关结果。Warther（1995）将资金流动分为预期流动和无法预期的流动，使用 Box - Jenkins 程序对预期流动进行估计。将无法预期的月度资金流动对同期股票市场收益进行回归，作者发现两者具有很强的相关性；但预期资金流动与股票市场收益之间却不存在这样的关系，因为预期收益已经包含在“同期之前”的贴现收益中，所以结论与有效市场理论一致。Warther（1995）也对反馈假设——颠倒回归方程中的因果关系方向，资金流动至少部分地由股票市场收益来决定——进行了检验，然而，作者没有为反馈假设找到经验证据。Remolona、Kleiman 和 Gruenstein（1997）也对“总资金流动 - 总的股票市场收益”之间的因果关系方向进行了检验，目的是为了确定在下跌的股票市场中，正反馈理论是否导致了股票价格的自持续（self - sustaining）下降。然而，他们的分析认为在 1986—1996 年期间短期股票市场收益对共同基金资金流动的影响非常微弱。

共同基金经理

同样重要的事情是注意如下事实，这里回顾的共同基金绩效研究文献都是将基金作为一个实体，而不是针对基金经理进行考察。在基金业发展历史中，基金经理很有可能遭到更替。因此，与考察基金绩效与基金特征（如投资目标、费用等）之间关系同样重要的是对基金绩效与横截面上基金经理特征之间的关系进行考察。Chevalier 和 Ellison（1999）对这一问题进行了研究，对基金绩效与基金经理特征之间的关系进行了考察，这些特征包括年龄、就读本科学校的平均入学成绩以及是否持有 MBA 文凭等。作者以 1988—1994 年期间独立管理某只基金的 492 位基金经理作为样本，结果发现，与没有 MBA 文凭的基金经理相比，持有 MBA 文凭的基金经理有 63 个基点的更高年平均收益，但持有 MBA 文凭的人似乎更喜欢投资于那些系统风险更高的股票，如果对此进行控制的话，是否拥有 MBA 文凭与基金绩效之间几乎没有关系。研究还发现，年轻基金经理的业绩稍好于年龄大的基金经理（对基金的其他风险特征进行控制），这可能是因为，处于职业生涯初期的基金经理面临更大被解雇的危险，从而他们更加努力地工作。最后，基金经理所就读本科学校的平均入学成绩与绩效之间有最为稳健的正相关关系，尽管基金绩效的这种差异可能与风险和费用等特征有关。Chevalier 和 Ellison（1999）认为，学生本科就读平均入学成绩更高的学校是学生更有天分、受到更好教育、拥有专业系统知识等的表征。

择时交易能力

为了对共同基金经理的择时交易能力进行校验，最理想的途径是获取实际的资产份额数据，考察基金经理在市场上涨之前是否增加了仓位。如果无法获得这样的数据，我们可以使用 Treynor 和 Mazuy（1996）或 Merton 和 Henriksson（1981）回归中所使用的条件变量。Treynor - Mazuy 的模型假定时变（time - varying）β 是基金经理对市场收益条件预测值的线性函数，即

$$\beta_t = \bar{\beta} + \lambda\ (\hat{R}_{m,t+1} - ER_m)$$

$\hat{R}_{m,t+1}$ 在这里表示基金经理在 t 时刻对市场收益的条件预测。Merton - Henriksson 模型所假设的关系与此类似，但 β 依赖于市场收益预测的符号，即 $\beta_t = \bar{\beta} + \lambda$（max [0，$\hat{R}^{+}_{m,t+1}$]）。

对择时交易能力的线性参数化处理要受到严格限制（Goetzmann、Ingersoll 和 Ivokvivh（2000）），这种处理无法对基金经理在择时交易时所获取的信息质量进行分离，也无法就基金

经理对这些信息的“进取性”反应进行考察。Jiang（2003）提出了对基金经理择时交易能力进行检验的非参数方法，这种方法对基金经理择时交易能力的质量进行甄别，方法不赖于线性因子模型。方法背后的直觉是，如果基金经理认为市场在下一个时期要上涨，那么他会把 β 的值调高。对任意两个不重叠的时期 t_1 和 t_2 而言，成功的择时交易意味着统计量 x 大于零：

$$x = 2 \cdot \text{prob}\ (\beta_{t_1} > \beta_{t_2} \mid R_{m,t_1+1} > R_{m,t_2+1})\ - 1$$

如果择时交易能力不存在，那么 β 与市场收益无关，从而 prob（·）$=1/2$，$x=0$（$x<0$ 意味着所出现的情形与择时交易情形相反）。现在考虑共同基金 i 的三元收益数组 $\{R_{i,t_1}, R_{i,t_2}, R_{i,t_3}\}$，且 $R_{m,t_3} > R_{m,t_2} > R_{m,t_1}$。具有卓越择时交易能力（不考虑他的“进取性程度”）的基金经理在 t_2 到 t_3 时期中的 β 值大于在 t_1 到 t_2 时期中的 β 值，两个时期的 β 可以分别表示为：

$$\beta_{12} = \frac{R_{i,t_2} - R_{i,t_1}}{R_{m,t_2} - R_{m,t_1}}$$

$$\beta_{22} = \frac{R_{i,t_3} - R_{i,t_2}}{R_{m,t_3} - R_{m,t_2}}$$

因此，可以将 x 类似地表示为：

$$\hat{\theta}_n = \binom{n}{3}^{-1} \sum_{\omega} \text{sing}(\beta_{23} > \beta_{12})$$

ω 取三个值，分别为 $\omega \in R_{m,t_1} < R_{m,t_2} < R_{m,t_3}$，sign（·）在{正，负，零}情形下分别取值{1，-1，0}。Jiang（2003）使用的是1980—1999年期间美国主动型共同基金的月度收益数据（样本包括1827只存活的基金和110只已经退市的基金，不包括指数追踪型基金）。某些 $\hat{\theta} > 0$ 的基金是因为运气好，但对具有择时交易能力的前5%的基金而言，$\hat{\theta} > 8.4\%$。因此，我们需要考察的是，对所有基金而言，在 $\theta = 0$ 的假设下，$\hat{\theta} = 8.4\%$ 是否位于5%的右尾之中。为了对此进行验证，假定服从正态分布，标准差取自基金的横截面数据。Jiang（2003）的研究发现，正态分布随机占优于 $\hat{\theta}$ 的经验分布（在 $\theta = 0$ 的零假设），基金经理不具有择时交易能力。因此，投资者最好是选择被动型指数基金，而不应在众多主动型基金中随机地进行选择。Jiang 还发现，择时交易的绩效与基金成立时间和管理层任期正相关，与基金规模负相关，但这些效应在统计上不显著，也没有经济意义。

与对英国共同基金的众多研究一样，得出美国共同基金绩效与绩效持续性研究的总体结论是一件非常困难的事情，因为不同文献的研究方法和样本期不尽相同，无法直接进行比较。然而，上述文献综述意味着，尽管基金经理具有卓越选股能力的证据不充分，但基金绩效在有些时候要高于风险调整后的基准投资组合收益，但这种绩效不会持续太久。与英国共同基金研究结论一致的是，绩效差的基金的持续性比绩效好的基金的持续性更为长久。

对冲基金

与共同基金一样，对冲基金也是将众多资金合在一起进行投资，但投资的限制要更少一些。市场上存在多种风格的对冲基金，比如说长短仓型、卖空型、多头杠杆型、新兴市场型、廉价证券型、全球宏观经济型和相对价值套利型。这些策略的风险都很大，所以术语“对冲基金”尤其具有误导性。对冲基金对最小投资额度的要求也比较高，每年的年费超过5%，而基金所有者获取的利润高达20%～25%。全球90%的对冲基金经理在美国，9%在欧洲，总共6000多家对冲基金管理着大约4000亿美元的资产。然而，许多对冲基金的规模相对较小，80%的对冲基金规模不到1亿美元。

由于数据缺乏，现有文献对对冲基金绩效的研究不是很多。早期研究认为对冲基金的耗损率（rate of attrition）特别高（Brown、Goetzmann 和 Ibbotson（1999）），一些研究发现对冲基金具有正的 α 值，从而风险调整后的绩效较好，绩效具有持续性（Agarwal 和 Naik（2000））。对对冲基金的绩效而言，投资风格的效应大于基金经理的能力差异（Brown、Goetzmann 和

Ibbotson（1999），Brown、Goetzmann 和 Park（2001））。Fung 和 Hsieh（1997）发现，对冲基金并不必然带来较高的风险调整后的收益，但它的收益与其他证券之间相关性较低，如果在指数追踪基金中纳入 10% ~20% 的对冲基金，那么组合的收益风险权衡关系会有所改进。

因为对冲基金具有多种投资风格，所以很难用一个线性因子模型来刻画它的收益，通常的做法是在 Carhart 四因子模型中增加一些收益变量，如世界股票指数的收益、各种债券指数的收益（包括新兴市场债券和公司债券的收益），甚至可以是商品指数的收益。尽管如此，这些“因子”是否能充分地刻画对冲基金的收益仍然存在很大的争议，因为对冲基金投资中包含了大量的期权，而期权偿付的非线性通常很高。

Capocci 和 Hubner（2004）使用多因子模型对包含2700 只对冲基金（包括801 只已经“死去”的基金）的大型数据库进行了考察，对 1994—2000 年（那些现在已经死去的基金当时尚在运作中）牛市行情中对冲基金的月度收益进行了研究。他们将对冲基金按照投资风格分为 20 多种类型（如长短仓型、小市值型等等），总体来看，在所有类型对冲基金中，约 25% ~30% 的基金的 α 值显著为正，5% ~10% 的基金的 α 值为负，剩下的大部分基金（即 60%）的 α 值为 0。对冲基金的市场 β 值小于共同基金的市场 β 值（约 0.3 ~0.6），SMB 因子的系数在统计上显著。对一部分对冲基金而言，新兴市场债券收益的系数也很显著，但只有三分之一对冲基金的 HML 因子的系数显著，约 15% 的对冲基金具有显著的动量因子。这些多因子回归模型的 R^2 基本上落在 0.65 ~0.95 的范围内。因此，大多数对冲基金都持有小市值股票，小部分对冲基金投资于新兴市场债券和动量型股票。

不幸的是，只有很少部分对冲基金（即长短仓基金、可转换套利基金、未分类的基金）在 1994—2000 年期间的正 α 值在一些子样本期中仍然为正。根据过去一年的收益（一月至次年一月）将所有对冲基金进行排序并分为十组，持有一年后再对投资组合进行调整（等权重），在这十个投资组合收益的时间序列中，只有两个（第 7 个和第 9 个）的 α 值在统计上显著，这意味着绝大多数投资组合的绩效不存在持续性。总的来说，一部分对冲基金在 1994—2000 年期间获得了正的异常收益，但是，“挑选赢利的股票”仍然很困难，绩效持续性看上去相当微弱。

9.3 共同基金之“星”

假设一些共同基金和对冲基金获得了异常收益，由此产生的问题是，这种绩效是否是选股能力的反映？是否是应归因于某个基金经理的好运气？直到现在，还没有文献对基金绩效中运气的作用直接进行建模。一些持续性研究通过将基金的排序期和评价期分离开来的办法来对运气进行控制。首先，如果用来对异常绩效进行排序的模型在两个时期中的设定都是错误的，那么这种方法就存在一定的缺陷。其次，递归投资组合的构造方法没有告诉我们，在各个十分之一组合或五分之一组合中，单只基金的绩效是否存在持续性，单只基金的绩效是否随着时间而发生改变。

正如我们已经看到的那样，大量证据表明，美国和英国共同基金的年平均收益比基准收益约低 0.5% ~1.5%（如标准普尔 500 指数，或其他类型的基准收益）。然而，也有一些基金经理（如成长型基金，Chen、Jegadeesh 和 Wermers（2000），Wermers（2000））确实获得了比基准收益更好的绩效。我们能确定所选择的基准是合理的吗？一些基金在多年中的“优良绩效”是否仅仅是因为“运气”？

在对美国共同基金绩效进行评估的文献中，这一假设会对业绩产生负面影响 Kosowski、Timmermann、White 和 Wermers（2003）是第一篇对运气进行控制的文章。面对数目如此庞大的基金样本，他们选择了 1700 只共同基金来进行研究，人们期望出现的一个结果是，一些基金的异常绩效仅仅是因为运气。在确定运气所导致的绩效边界时就会产生这样的问题。

正如我们已经看到的那样，为了对共同基金的平均绩效进行检验，对每只基金而言（$t=1, 2, \cdots, T$），我们需要一个预期收益模型。这里为了便于说明问题，我们选用“无条件”CAPM 模型（即参数为常数）：

$$(R_{it}-r_t) = \alpha_i + \beta_i (R_{mt}-r_t) + \varepsilon_{it} \tag{10}$$

对任意基金而言，无异常绩效的零假设为 H_0：$\alpha_i = 0$。为了对基金的平均绩效进行检验，我们针对每一只基金对（10）进行回归。如果总共有 N 只基金，那么 N 只基金的平均 α 和 t_α 值分别为：

$$\alpha_{av} = N^{-1}\sum_{i=1}^{N} \hat{\alpha}_i \tag{11a}$$

$$t_{av} = N^{-1}\sum_{i=1}^{N} t_{\alpha i} \tag{11b}$$

我们已经注意到，对无条件 CAPM 模型而言，我们感兴趣的替代模型是条件模型，条件模型中的 α_i 和 β_i 都具有时变性。如果 z_t 是 t 时刻已知的信息集合，那么条件模型为：

$$(R_{it}-r_t) = \alpha_i(\mathbf{z}_{t-1}) + \beta_i(\mathbf{z}_{t-1})[R_{bt}-r_t] + \varepsilon_{it} \tag{12}$$

$\alpha_i(z_t)$ 和 $\beta_i(z_t)$ 都是线性函数：

$$\alpha_i(z_t) = \alpha_{0i} + \alpha'\mathbf{z}_t \tag{13a}$$

$$\beta_i(z_t) = \beta_{0i} + \beta'\mathbf{z}_t \tag{13b}$$

变量$\mathbf{z}_t$ 对因子载荷和因子风险溢价中可预测变化进行了控制。向量$\mathbf{z}_t$ 包含的变量有：红利收益、长短期利率差、短期国债利率和表示一月效应的虚拟变量。现在，异常绩效可以用 $\alpha_{0i}\neq 0$ 来表示。

美国共同基金之“星”？

Kosowski、Timmermann、White 和 Wermers（2003）使用美国 1700 只共同基金 1962—1994 年期间的月度收益率数据来进行研究，结果发现，$\alpha_{av}=0.003\%$（每月），$t_{av}=-0.004$，这意味着，平均而言这些基金的绩效没有超过市场基准。但是，1700 只基金的 α_i 值在横截面上的标准差高达 0.3%（每月），这意味着样本内一些基金的绩效确实很好（一些基金的绩效确实很差）。他们据此提出的问题是，绩效特别好的“赢家”（和绩效特别差的“输家”）是否仅仅是因为运气的好坏？

为了解决这个问题，他们使用了自助法，假设方程的残差序列不服从正态分布。自助法的优点在于，能为绩效的统计推断提供一种非参数方法，绩效回归方程的残差一般不会服从正态分布，这使得基于标准 t 检验和 F 检验的统计推断变得不可靠。

我们在此说明如何对基金使用自助法。首先，我们使用真实数据可以估计出某个特殊基金的 β_i 值（比如说，0.9）和残差序列 $\hat{e}_{it}$。在自助法的使用过程中，我们让 $\beta_i=0.9$ 和序列（$R_{mt}-r_t$）保持不变。现在，对这只基金而言，我们以相同的概率随机地选择 $\hat{e}_{it}$来进行替换，生成样本容量为 T 的（$R_{it}-r_t$）序列，在绩效零的零假设下（也就是说 $\alpha_i=0$）：

$$R_{it}-r_t = 0 + \beta_i (R_m - r)_t + \hat{e}_{it} \tag{14}$$

使用（$R_{it}-r_t$）的拟合数据，根据（10），我们可以计算（$\hat{\alpha}_i$，$\hat{\beta}_i$）的最小二乘估计值和 t 统计量 t_{α_i}的值（使用 GMM 标准差）。因为我们对 $\hat{e}_{it}$进行了“新”的随机抽取，所以采用自助法的估计结果与使用真实数据的估计结果不同。例如，如果随机抽取由于偶然因素出现的大量负值，那么 $\hat{\alpha}_i$ 的自助法估计值为负。

我们对所有基金 $i=1, 2, \cdots, N$，重复上述程序就得到 α_i 和 t_{α_i}的 N 组值。为了便于说明问题，假设我们感兴趣的问题是，基金的最好绩效是否是由运气所导致的。我们可以从 α_i 和 t_{α_i}的 N 组值中选择最大的 α_i^{boot}（max）和 $t_{\alpha_i}^{boot}$（max）（即所有基金中绩效最好的基金）。我们就完成了第一轮的自助法计算（*boot* = 1），接下来重复上述程序 1000 次，我们就得到 1000 组

α_i^{boot}（max）和 $t_{\alpha_i}^{boot}$（max），由这1000组值就可以得到绩效最好基金的自助法经验分布（在绩效为零的零假设下，也就是说，$\alpha_i=0$）。

重复上述程序就可以得到任意绩效基金的（$\hat{\alpha}_i$，$\hat{\beta}_i$）的自助法分布（例如，在基金绩效排名中，位于2%，3%，…，最差和最好的基金）。事实上，t 统计量的自助法分布比 α 本身的自助法分布有着更好的性质（即它是一个“枢轴统计量”），直观原因是，t 统计量要用 α 值的标准差来进行调整，因此，t 统计量自助法分布的分散程度比 α 本身自助法分布要小得多。Kosowski 等报告了每一绩效水平所对应 α 的自助法估计值之间的差异，结果发现，在分布的极端尾部中，差异变得非常大。

我们来考察绩效排名处于第五位置的基金的自助法分布。Kosowski、Timmermann、White 和 Wermers（2003）使用美国共同基金真实数据估计的 $\hat{\alpha}$ 值为每月1%，相应（GMM）t 统计量为4.0。对所有基金而言，自助法的零假设为 $\alpha_i=0$。对（绩效排名位于第五位置的基金的）t 统计量自助法分布而言，在 $\alpha_i=0$ 的零假设下 p 值小于0.001。因此，绩效排名处于第五位置的基金由于运气而获得每月1%（α）绩效的概率非常小。在对运气因素进行控制之后（对横截面上的所基金而言），我们可以确信，绩效排名第五的基金每月1%左右的绩效应归因于杰出的选股能力。

我们用一种稍稍不同的方式来对此进行阐述。Kosowski、Timmermann、White 和 Wermers（2003）根据 $\alpha_i=0$ 时的自助法拟合数据发现，人们所期望的结果是应该有9只基金仅仅是因为运气才获得每月 $\alpha_i \geq 0.5\%$（即每年6%）的绩效。但是，在真实数据中有60只基金获得了这样的绩效，因此，一些基金的基金经理确实有着优秀的选股能力。主动型基金的管理具有较强的存活能力。他们对不同“风格”的基金重复上述分析时发现这些基金中的“明星”大部分是“成长型基金”。

Kosowski、Timmermann、White 和 Wermers（2003）也发现，一般而言，位于分布非常靠左区域中基金（例如，绩效最差的5%~10%基金，或绩效最差的1%~10%的基金）的绩效非常差，低劣绩效不能归因于运气不好，这些基金的选股技能确实很差。

值得注意的是，这些结果说明某些基金卓越的选股技能在很多月份中确实能获得比平均绩效高的收益。如果这种绩效在未来具有一定的持续性，那么你现在就应该选择绩效最好的基金以便在将来击败市场基准。当然，就现实情形而言，存在太多的变数（如基金选股技能的变化、基金管理层的变更等等）以至于上述情形不会发生。

Kosowski、Timmermann、White 和 Wermers（2003）使用 Carhart（1997）的程序来考察基金绩效在一年中的持续性。在每年的一月份，根据 Carhart 四因子模型和过去三年数据所获得的 α 值来对所有基金进行排序和分组，就每一个组合而言使用等权重的方法来构造投资组合，投资组合的持有期为一年，一年后再重复上述程序并对组合进行调整（比方说绩效最好的十分之基金组合中有一部分基金每年都会发生变化），然后使用1979—1993年的真实数据来估计每一组（如绩效最好的10%基金组合）的 α 值。在 $\alpha_i=0$ 的（横截面）约束下，使用相同程序就可以得到每一个组的自助法分布。根据自助法的结果，对绩效最好的10%的基金（$\alpha_i=0$）使用等权重方法来构造一个投资组合，每年对组合进行一次调整，由此可以得到绩效最好基金组合分布的 $\hat{\alpha}_i$（和 t_{α_i}）值。他们对所有基金进行研究后发现，只有绩效最好（前10%）那一组基金在一年中才具有持续性，α 的值达到每年1.6%。对进取型基金而言，绩效最好的三个组合在一年中都呈现出一定的持续性，而对成长型基金而言，只有绩效最好的那一组才具有持续性，其他投资风格的基金（如成长收益型、平衡收益型）则没有持续性。

这种“优良绩效”依赖于具体的模型，CAPM 模型可能不是最为恰当的模型。Kosowski、Timmermann、White 和 Wermers（2003）使用多种模型来对基金绩效进行考察（如 Fama - French 三因子模型、Carhart 包含“动量因子”的四因子模型、几个条件 β 模型），结果发现，上述结论基本上没有什么变化。尽管你只有0.5%的几率选择到能战胜市场的“明星基金”，获得每月1%（每年12%）的收益仍然是运气使然。

Kosowski、Timmermann、White 和 Wermers（2003）在这一领域的研究最为深入，他们使用包括 *CAPM* 在内的多种模型来考察绩效最好和最差的 1%、5%、10% 等基金的绩效，同时将幸存偏差控制到最小（例如，保留那些只在样本期中某一段时间存活的基金）。总的来说，他们得出的结论似乎很稳健：1700 只基金中绩效最好的 5～10 只基金确实能战胜市场，原因是选股能力强，而绩效最差的 5～10 只基金的选股能力的确太差。

上述结论似乎存在一个矛盾：一方面，市场中存在如此多的“主动型”共同基金，另一方面，用 Jensen 度量的这些基金的绩效却很差。当然，有许多基金的绩效比被动型投资组合好，也有一部分基金的绩效比被动型投资组合差。困难在于找出谁是“赢家”，考察“赢家”中的一个较小部分在将来是否仍能成为“赢家”。这不可能是“免费午餐”。

自助法中的一些问题

自助法使用者在应用中需要注意一些问题。我们可以不用 α 来对基金绩效进行度量，但可以在无异常绩效 H_0: $\alpha_i=0$ 的零假设下对“α 的 t 统计量”自助法分布进行估计，然后将这个分布与所观察到的 t_α 进行比较。使用 t 统计量还有一个好处：对那些观测值较少的基金而言，估计的方差会更大一些，估计的精度要更低一些，也就更有可能产生异常的 α 值。由此导致的一个风险是，这些基金在实际 α 分布和使用自助法所估计的 α 分布的极端尾部中具有不对称分布。t 统计量通过 α 值的尺度变换对估计精度（即标准误差）进行了修正。相对于绩效分布中同一百分比位置的 α 自助法分布而言，极端值的自助法 t 统计量的分布由的大方差所导致的问题可能更少一些。在这种类型的研究中，α 及其 t 统计量的自助分布按照通常的方式来进行构造。（所有 t 统计量的构造都是以用标准误差调整后的 Newey－West 异方差和自相关性为基础。）

为了进一步提高绩效估计的精度，我们可以将那些观测值最少的基金纳入到分析之中。因为观察值不是太多，所以估计结果在不同的样本中有很可能发生很大的变化，这会增加自助法分布尾部的宽度，使得结果偏离如下结论：基金绩效不是人们所预期的那样由偶然因素所导致。例如，分析中将 60 个观测值设定为合适基金样本的最小观测值个数。然而，这种方法的一个不足之处在于施加了某种形式的幸存偏好，因为分析过程要对所考察的基金进行限制，比如选择那些选股能力强的基金，选择那些存活时期超过五年的基金。为了考察这个问题的意义，我们可以通过为最小观测值个数设定多个数值的方法来检验自助法结果的敏感性。

我们可以使用各种自助法重新选样技巧，这些技巧可以只生成新的残差样本，也可以同时生成残差和因子的样本（从而保留真实数据中所观察到的时间序列配对关系），也可以独立地生成残差和因子的样本。

也有可能出现的情形是，我们对基金残差在横截面上的独立性——对基金 i 和 j 而言有 Cov $(\varepsilon_i,\ \varepsilon_j)\neq 0$——表示怀疑。绩效模型的错误设定会导致这一问题。为了在自助法的使用过程中保留这种相关性信息，在 1000 次自助模拟的每一轮计算中都需要从所有基金的时间序列观察值中进行随机抽取。然而，使用这种方法存在一个困难，因为样本中的基金并不同时存在（也就是说，有的基金在其他基金进入样本之前就从样本中删掉了）。要使自助法分析结果与同期基金发展的历史相吻合就必须对要分析的基金样本进行严格限制，由此导致的问题是结论不一定可靠。消除这一问题的方法是使用各种均衡绩效模型，将错误设定和基金横截面上的相关性限制在一定范围之内。

在自助法的使用过程中，我们在每一次随机生成样本残差时都假设残差独立同分布。然而，如果共同基金样本呈现出（比方说）一阶序列相关性，那么对不同长度块中的残差使用自助法也可以对这种相关性进行考察。

我们也感兴趣的问题是基金经理的选股能力是否与基金的投资目标相关，对不同投资目标的基金（比如说股票收益型、股票成长型、普通股票型和小公司型）分别使用自助法就可以对这问题进行考察。为了对基金绩效进行全面的研究，为了对结果的稳健性进行检验，我们

应该对各种绩效度量模型（也就是说，无条件的单因子模型和多因子模型，以因子载荷和为条件的单因子模型和多因子模型）使用自助法。Kosowski、Timmermann、White 和 Wermers（2003）使用了各种自助法技巧，总的来说，他们的发现是所有这些方法得出的结论都相同。这意味着驱动绩效分布尾部中自助法结果的残差是非正态的。

英国共同基金之“星”?

现在，我们来描述英国共同基金的一些初步研究结果，内容来自于 Cuthbertson、Nitzsche 和 O'Sullivan（2004）。为了寻找共同基金之“星”（也就是说，即使对可能的“好运气”进行考虑后，绩效在所有基金中都是非常好的那些基金），他们对共同基金使用了多种自助法技巧。

数据

共同基金数据库由 1596 只单位信托基金和开放性投资公司（OEICs）组成。单位信托基金将众多投资汇集在一起，使投资者分享充分分散化证券投资组合所带来的收益。然而，根据信托契约的宗旨，单位信托基金的投资目标也常常不同。从某种意义上说，单位信托基金是“开放式”共同基金，因为投资者随时都能以现行市场价格购买新的基金份额，对基金份额的需求不会使单位价格增加。单位信托基金只能在投资者和信托经理之间进行交易，没有二级市场。单位信托基金与投资信托基金有区别，后者常用术语“封闭式”基金来描述。投资信托基金尽管也是将众多投资汇集在一起，它在股票交易所的报价是针对所有权而言的，也按照证券的方式进行交易。与普通的股份公司一样，信托投资基金的份额数是固定的。与单位信托基金不同的是，对投资信托基金的需求会提升每单位基金的价格。因此，当价格高于（低于）资产的内在价值时，投资信托基金有可能进行溢价（折价）交易，溢价（比如说）有可能反映了投资者对投资信托基金经理投资技能的认同。OEICs 按照公司的形式进行组建，投资者可以购买它的股份，但发行的股份数会根据需求的变化而变化，所以人们用术语“开放式”来进行描述。这意味着价格总是反映资产的内在价值，与投资信托基金不同的是，它的价格不受市场对 OEICs 情绪的影响。因此，OEICs 和单位信托基金的风险特征较投资信托基金更为明显。

共同基金的数据来自于标准普尔基金评鉴公司（Standard & Poor's Micropal），收益使用的是 1975 年 4 月至 2002 年 12 月的月度数据，基金的选取限制于投资英国股票的基金，划分的依据是至少 80% 的资金投资于英国的股票，划分标准来自于投资管理协会（IMA，Investment Management Association），其前身是单位信托基金与投资信托基金协会（AUTIF，Association of Unit Trusts and Investment Funds）。

在 1596 只基金中，有 334 只基金属于附属基金（second units），这是因为在基金管理公司发行一只新基金时，它实际上发行了几种类似的基金。一些不太成功的基金后来就被合并到较为成功的基金之中，因此附属基金都是不太成功的基金。附属基金仍然被留在数据库中，但合并后所报告的收益是较为成功基金（合并方）的收益。合并前的收益非常相似。当某只基金被分拆成若干只相似基金时也会导致附属基金的产生，在这种情形中，附属基金在数据库的记录时期是整个样本期，但报告的是分拆前的收益。在绝大多数情形中，分拆发生在基金运行的后期，因此这些附属基金有着相对较少的“独立”收益观测值。无论附属基金是在哪种情形中产生，它们所提供的信息都非常少。这一研究将 334 只附属基金挑选出来，在大多数分析中将他们排除在外。此外，数据库中有 128 只基金属于市场追踪型基金。因为研究感兴趣的是考察共同基金相对于市场指数基准的绩效，所以指数追踪型基金也被排除在外。最终样本包括 1150 只非市场追踪型独立基金（即非附属基金）。

他们根据投资目标来对所有股票型基金进行分类。投资目标分为股票成长型（122 只基金）、股票收益型（253 只基金）、普通股票型（即收益型与成长型，591 只基金）和小公司型

(184 只基金)。IMA 基金分类系统认为，股票收益型基金是那些能取得超过金融时报全股指数收益 110% 的基金；股票成长型基金的首要目标是实现投资的增长；普通股票基金要求 80% 的资金投资于英国股票；小公司基金要求至少 80% 的资金投资于英国股票，在总股本中至少 10% 是市场资本。这些投资风格由基金公司自身进行确定，但最初的认定和随后的每月监督则由英国 IMA 进行。如果基金在运行中连续三个月没有遵循自己规定的投资目标，那么它们必须改变现状或更改投资目标。

为了对幸存偏好进行控制，数据库既包括存活基金（765 只），又包括非存活基金（358 只)。此外，也可以根据基金的运营地来对基金进行分类：英国本土基金（845 只）和海外基金（305 只）。海外基金主要集中在都柏林、卢森堡和英属海峡群岛。

收益率的计算采用“买价对买价”（bid price to bid price）的方法，所有收益都进行再投资。基金收益中的买卖价差包括初始佣金（通常为 5 ~6%)、印花税、交易费、基金管理费和标的证券竞买竞卖价差。因为研究采用的是“买价对买价”的计算方法，所以价格包含所有交易成本。基金对顾客收取的年费已经纳入到信托价格之中，Micropal 数据库无法将年费分离出来。因此，年费要从收益中扣除，这可能会产生一些担忧，因为这意味着在绩效模型的估计中，基金收益是扣除年费后的收益，而市场基准指数收益则是用毛收益进行计算的收益。这会影响对因子载荷的估计。然而，如果假设年费是相对固定的（固定的百分比年费就是这种情形)，那么因子载荷的估计不会受到任何影响，(固定的）年费可以简化绩效指标。(当然，实际操作中还应将年费分摊到每个月的观测值中去。) 从 1997 年开始，我们就可以得到总费率（TER）数据，总费率主要由年费构成。为了强调上述担忧，人们可以将 TER 纳入到净收益之中，在单位信托基金的整个运行期间，TER 似乎一直不变。然而，正如我们已经指出的那样，这一程序等价于将 TER 加到 α 的估计值上（每月要进行合适的调整)。

无风险利率选用的是英国一月期国库券利率，市场指数使用的是金融时报全股指数。对规模效应而言，所使用的因子模拟组合 SMB_t 等于 Hoare Govett 小公司指数（HGSC）月度收益减去金融时报 100 指数（FTSE 100）的月度收益。HGSC 指数用于测度英国证券市场上资本规模最小的 10% 公司的绩效。

对价值溢价而言，所使用的因子模拟组合等于摩根斯坦利资本国际（MSCI）英国价值指数的月度收益减去 MSCI 英国增长指数的月度收益。为了构造这些指数，摩根斯坦利将英国股票指数中所有股票按照账面市值比进行排序。从账面市值比最高的股票开始，所有股票都归属于价值指数，直到市值达到英国国内股票市场总资本的 50% 为止。其他股票则归属于增长指数。MSCI 英国指数至少具有 60% 的市场代表性（这一比例最近提高到 85%)。设计这一指数的目的是为了广泛地代表各种类型的股票而不是仅仅代表市值最高的股票。

用于捕获一年期动量效应的因子模拟组合可以根据金融时报 100 指数来进行构造。在每一个月中，$PR1YR_t$ 等于过去 11 个月中收益最高 30% 股票的等权重加权平均收益（equal - weighted average return）减去收益最低 30% 股票的等权重加权平均收益。变量要滞后一个月。在条件因子载荷模型和条件绩效度量模型的估计中使用的工具变量如下。无风险利率是英国一月期国库券的赎回收益率，市场因子的红利收益率是金融时报全股指数的红利收益率，期限结构的斜率等于英国 10 年期金边债券（UK10 - year gilt）的收益率减去一月期国库券收益率。

模型选择

CAPM 模型、Fama - French 三因子模型和 Carhart 四因子模型、Treynor - Mazuy 模型和 Merton - Henriksson 市场时变模型等无条件模型的估计结果大致相同。比如说，CAPM 模型表明 α 的横截面平均值大约为每月 -0.03%，这意味着平均来看共同基金经理的绩效每年比市场低 0.36%，但这一异常绩效在 5% 的显著性水平下不显著（所使用的统计量是用标准误差对 Newey - West 异方差与自相关性进行调整后的各个基金所对应绝对值的平均)。Fama - French 模型和 Carhart 因子模型所得到的绩效结果类似。

根据因子载荷，所有无条件模型的 t 统计量都表明市场风险因子（market - risk factor）和规模风险因子（size - risk factor）是股票收益横截面变化的显著决定因素。例如，在 Carhart 四因子模型中，市场风险因子所对应 t 统计量的横截面平均值约为 25，而规模风险因子所对应 t 统计量的横截面平均值约为 5.4。事实上，就“市场”和“规模”两个风险因子所对应 t 统计量而言，分别有 100% 和 80% 的共同基金在统计上显著。“价值”风险因子对基金收益的解释似乎没有显著的影响，就这个风险因子而言，只有 24% 的共同基金所对应 t 统计量是显著的。在 Carhart 四因子模型中，一年期动量因子相对来说不重要，只有 21% 基金的动量效应是收益的重要决定因素。

在条件 β、条件 α 和 β 模型中，与 α 及其分布有关的结果与无条件因子模型的结果非常相似。平均而言，几乎所有模型都统计显著地认为共同基金经理没有战胜市场。条件因子模型的设定允许对投资组合的敏感性进行动态调整，从宽泛意义上讲，这些设定体现的是部分基金经理的择时交易能力。上述参数检验为如下结果提供了证据：从整体来看，基金经理要么没有择时交易能力，要么做得不是很成功。（结果与检验英国单位信托基金择时交易能力的文献结果一致（参见 Fletcher（1995），Leger（1997）。）而参数检验难免要涉及一个联合假设，Jiang（2003）对美国股票型共同基金进行非参数检验却发现这些基金并没有很好的择时交易能力。

就关键的模型选择标准而言，研究使用的是 Schwartz 信息准则（Schwartz Information Criterion，SIC）对“拟合效果最好的”模型而言。Schwartz 信息准则试图减小模型残差的平方和。在 Schwartz 信息准则下，SIC 值最小的模型就是应该选择的模型。在各种无条件模型中，三因子 Fama - French 模型具有最小的 SIC 值（对所有共同基金的平均值而言），事实上，这个模型给出了最好的拟合效果，Kosowski、Timmermann、White 和 Wermers（2003）也报告了类似结果。

就各种模型设定而言，60% ~ 80% 的共同基金都拒绝正态性假设（使用 Jarque - Bera 检验）。这一结果极大地促进了自助法的使用，因为非正态的残差意味着估计值本身也呈非正态分布，这又反过来使得标准统计检验（如 t - 检验和 F - 检验）变得无效，标准统计检验要求误差服从正态分布。就共同基金的异常绩效而言，以往的研究根据检验来进行推断，众多基金残差的非正态性使我们对这些推断的可靠性产生了怀疑。这极大地促进了人们对绩效估计使用自助法，目的是为了确定共同基金业中是否存在显著的异常绩效（或拙劣绩效）。

对所有模型而言，超过 60% 基金的残差具有一阶序列相关性（使用 LM 检验统计量）。这意味着在使用自助法分析时对残差的随机抽取必须成对进行。为了使自助模拟法能尽可能地拟合原始数据生成过程，对序列相关性中的信息保留而言，上述技巧也是必需的。

英国共同基金的自助法分析

正如前面所注意到的那样，在“样本内”Schwartz 信息准则的基础上，研究发现无条件 Fama - French 三因子模型的表现“最好”。表 3 报告的是对所有共同基金样本（即所有投资目标的基金）使用无条件 Fama - French 三因子模型的自助法统计量结果。表中报告了“实际的”α 估计值和用 Newey - West 方法调整后的 t_α 值，表中的值是在无异常绩效零假设下（即对所有基金 i 而言，$\alpha_i = 0$）t_α 所对应的概率值。这一节中报告的所有自助法结果都是基于 1000 次模拟的结果。

表 3　　英国共同基金绩效

基金位置	无条件三因子模型		
	实际 α 值	实际 t_α 值	自助法 p 值
绩效好的基金			
最好的基金	0.7853	4.0234	0.056
第 2 好的基金	0.7239	3.3891	0.059
第 10 好的基金	0.5304	2.5448	0.022
第 15 好的基金	0.4782	2.4035	0.004
绩效差的基金			
第 15 差的基金	-0.5220	-3.6873	0.000
第 10 差的基金	-0.5899	-4.1187	0.000
第 2 差的基金	-0.7407	-5.1664	0.001
最差的基金	-0.9015	-7.4176	0.000

注：该表列出了在整个共同基金样本期中自助法统计量的结果，包括对所有投资目标的基金使用 Fama - French 三因子绩效模型所得到的结果。第一列是使用真实数据所估计的值，第二列是使用真实数据和用标准差对 Newey - West 异方差和自相关性进行调整后的统计量 α 值，第三列是在无异常绩效的零假设（α_i）下在 1000 次自助重新采样的基础上所得到的自助法 p 值。样本期间为 1975 年 4 月至 2002 年 12 月。

我们以绩效最好的基金为例。表 3 表明，按照无条件模型的 α 估计值进行排序，绩效最好的基金能实现每月 0.785% 的异常绩效，0.056 的自助法 p 值意味着在对所有基金进行 1000 次自助法模拟中（在零异常绩效的零假设下）所得到的 t_α 比“真实”数据中“绩效最好”基金的 t_α 值大的频率为 5.6%。换句话说，0.056 的 p 值意味着在 95% 的置信水平下可以认为最好的基金绩效（以 t_α 作为绩效度量指标）在统计上应归因于运气，部分基金经理具有卓越的选股能力并没有得到证据的支持。

接下来考察表 3 中绩效分布的整个右尾，证据表明，有些基金（恰好位于分布的极端右尾中，如绩效第 10 好的基金和绩效第 15 好的基金）的基金经理确实拥有出众的选股能力，但绝大多数“卓越”的基金绩效应归功于运气。这一结果与 Kosowski、Timmermann、White 和 Wermers（2003）的结果有些不同，他们寻找到了最好基金经理具有“选股能力”更强的证据。

在分布的左尾中，绩效最差基金所对应的 t_α 为 -7.4，自助法 p 值接近于零。也就是说，在对所有基金进行的 1000 次自助法模拟中（在 $H_0: \alpha_i = 0$ 的零假设下），t_α 比真实数据中所发现的 -7.4 还小的频率为零。这意味着，糟糕的基金绩效不是因为运气不好。表 3 所报告的另外一个明显事实是，上述结论在所有的左尾分布中都成立。在分布的负绩效分布一侧所得到的结论与 Kosowski、Timmermann、White 和 Wermers（2003）对美国研究的结论相似，他们也发现，绩效糟糕的基金确实是因为选股能力太差。

图 2 对 t_α 的自助分布与所观察到的 t_α 的分布进行了比较，前者完全是因为随机变化，后者是使用基金样本的“真实”数据所得到的值。该图描述了用无条件三因子模型对所有投资目标的基金所进行的核密度估计结果。从图 2 可以明显看出，t_α 真实绩效分布的左尾比自助法分布的左尾更厚，但右尾只在很小的范围内落在自助分布右尾的外面，这再一次证明确实有不少绩效差的基金存在，对位于分布右尾的基金而言，t_α 应主要归因于运气。

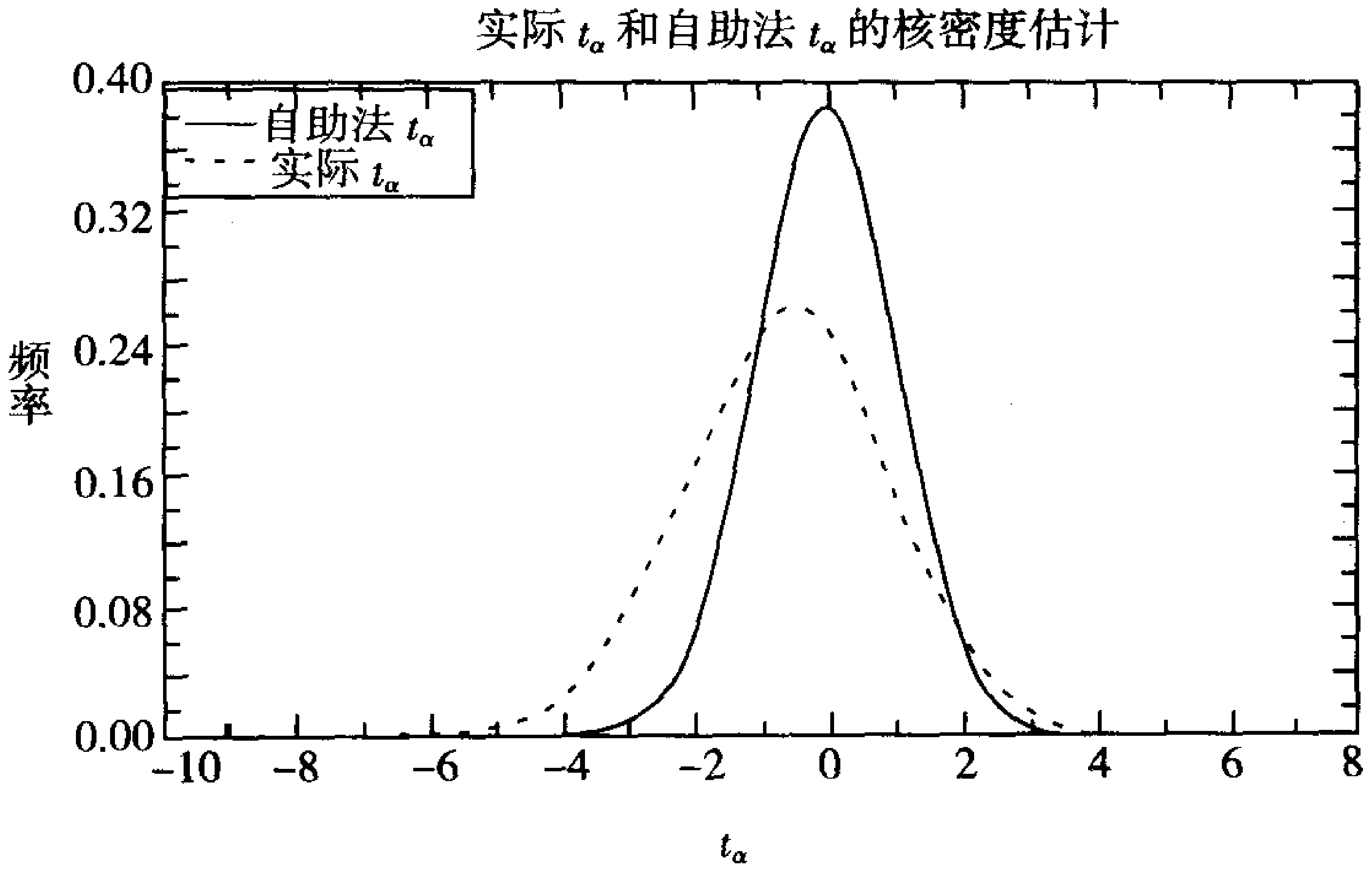


图2 实际 t_α 的核密度估计和自助法分布：英国所有共同基金（无条件三因子模型）

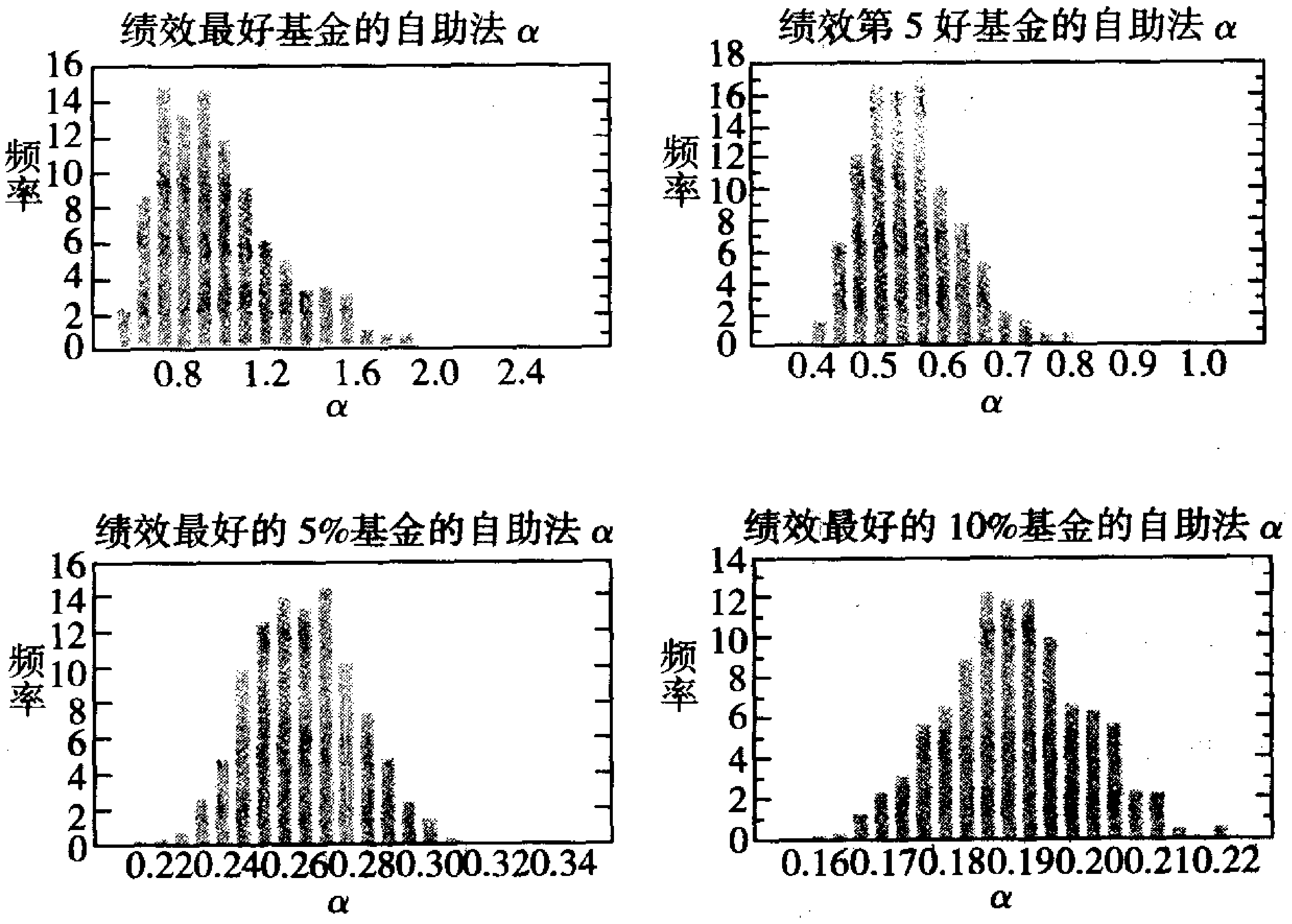


图3 α 自助法估计值的柱状图：英国所有的共同基金
（分布的上端：无条件 Fama - French 三因子模型）

图3和图4给出的是绩效分布中某些点所对应的 α 值的自助法分布。图3左上角的图形给出的是在无异常收益零假设下根据1000次自助法重采样本所得到的绩效最好值的分布，而右上角的图形给出的是绩效第5好 α 值的1000次自助法分布，其余类似。从图3中的四个子图可以明显地看出，绩效最好基金的自助法分布具有高度非正态性，具有相对较大的方差，当我们更加靠近绩效分布的中心时，分布与正态分布更为接近，方差也更小一些。图形所没有显示出来的是，这一结果与如下事实非常接近：在绩效“最好”基金的回归分析中，与更靠近分布中心的基金相比，残差具有更大的方差和更严重的非正态性。正是绩效最好基金回归分析的残差具有大的方差导致了自助过程中 α 值有着更大的分散性。对绩效分布较小的尾段而言，图4几乎是图3的反面，对绩效特别差的基金回归分析而言，残差具有很大的方差且不服从正态分布，这也会导致自助法 α 有着较大的分散性。

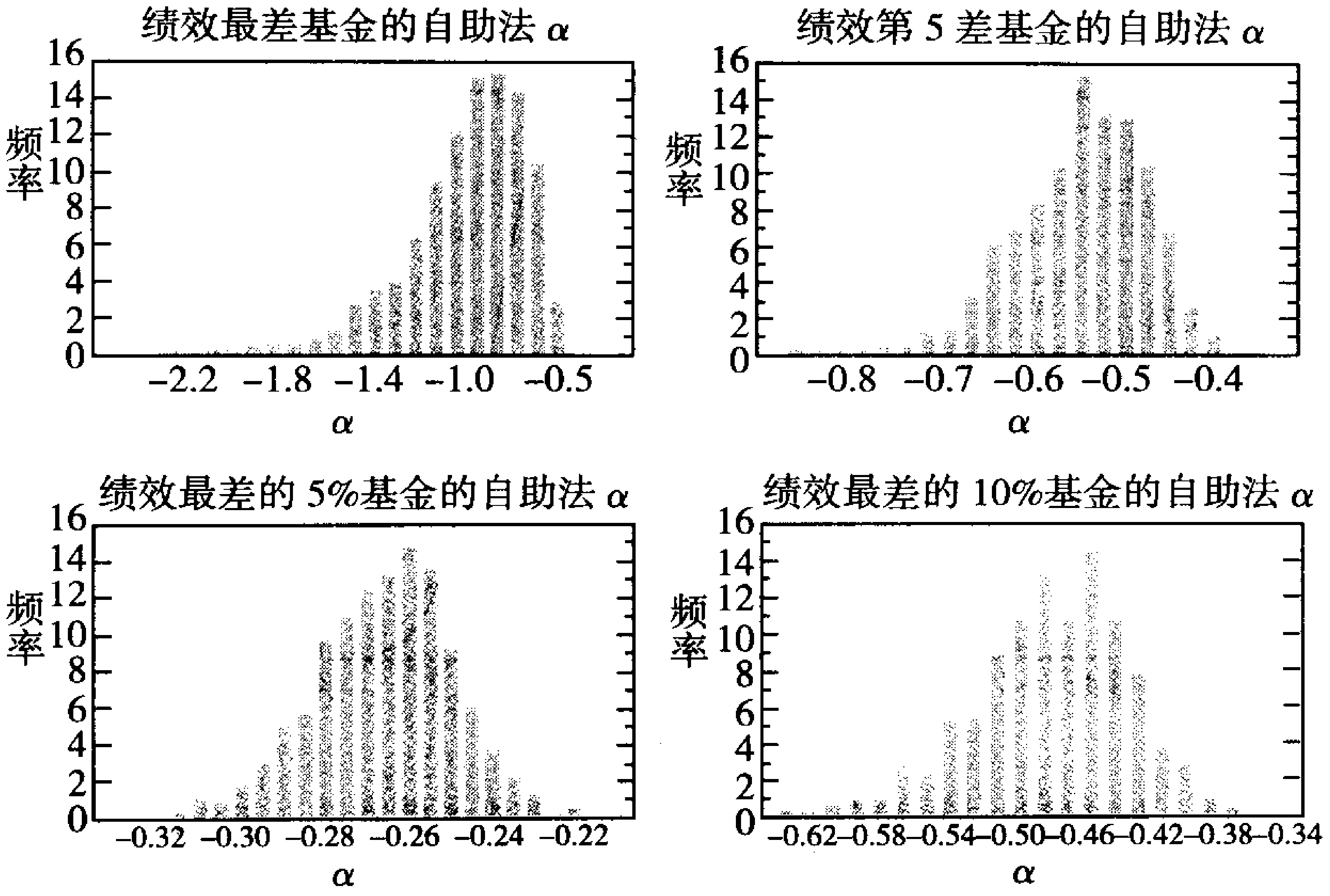


图 4　α 自助法估计值的柱状图：英国所有的共同基金
（分布的下端：无条件 Fama - French 三因子模型）

作为一种替代解释，自助法可用来估计在基金样本中进行随机选择后有多少基金能实现某个给定的 α 绩效水平。我们可以将这个数值与实际实现这一 α 绩效水平的基金数目进行比较。例如，根据概率在自助法中应有 364 只基金具有正的 α 值，但实际上我们只观察到 240 只基金具有正的绩效。

绩效与投资风格

对投资者而言还有一个感兴趣的问题：基金经理的选股才能是否与基金的投资目标有关系？在以往的共同基金绩效文献和基金绩效持续性文献中，尤其是在对美国基金业进行研究的文献中，有证据表明，投资成长性股票的共同基金往往是绩效最好的基金（Chen，Jegadeesh 和 Wermers（2000））。

在这一节的研究中我们发现，在根据三因子无条件 α 模型进行排序的绩效最好的十只基金中，三只是成长型基金，四只是普通股票基金，还有三只是小公司股票基金，没有收益型基金。在这十只基金中，只有五只存活到 2002 年 12 月。就 α 绩效的另一端而言，在绩效最差的十只基金中，五只是普通股票基金，两只是收益型基金，三只是小公司股票基金，没有成长型基金。在绩效谱的顶端，这里的结果支持美国共同基金的研究结论：成长型股票基金倾向于在绩效最好的基金中出现，但不会在绩效非常差的基金中出现。

为了更为严格地强调不同投资风格共同基金的相对绩效问题，这一节的研究分别对每一种投资风格或目标的基金使用自助法程序。基金投资风格由基金自身的发行公告来界定，但一开始的认证和之后的每月监督则由英国的 IMA 来进行。在每个基金类别中对基金经理的绩效和能力进行考察的一个优势是，在每一种情形中，更为同质的风险群体有助于控制可能的、未知的横截面风险特征，这些风险特征有可能是收益均衡模型的错误设定所导致的。

图 5 至图 8 对不同投资风格的基金绩效进行了比较。这些图分别给出的是股票收益型基金、股票成长型基金、普通股票型基金和小公司股票型基金实际 t_α 和自助法 t_α 分布的核密度估计结果。对这些图形进行比较就会发现：对每一组而言，实际绩效分布（用虚线表示）的左尾在很宽的范围内都很相似，糟糕的基金绩效并不是因为运气不好。只有成长型经济和收益型基金的实际绩效分布的右尾位于零绩效自助法分布的右侧，即使是这样，也没有充分的证据来证明这是源自能力而非“好的运气”。

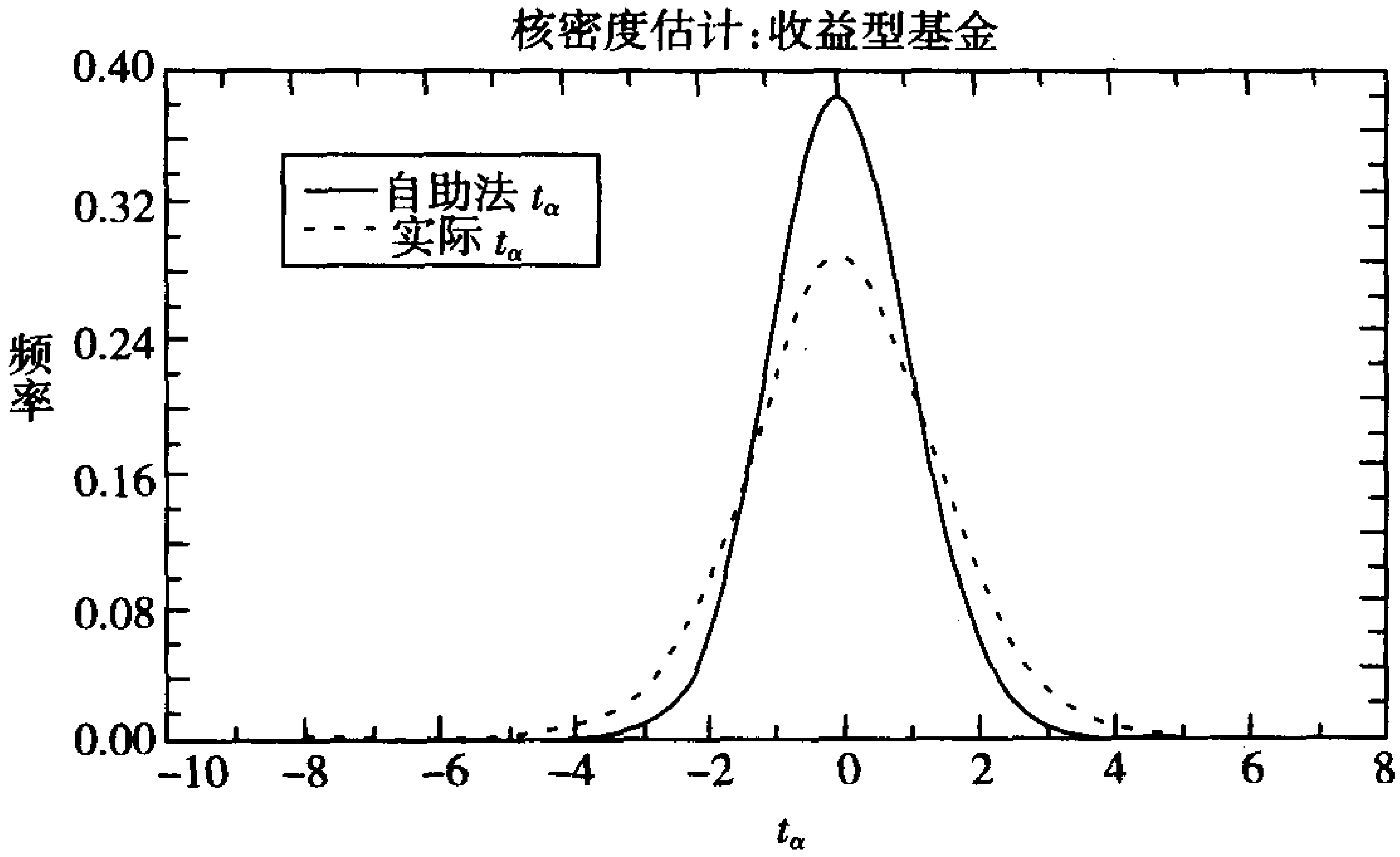


图5 实际 t_α 的核密度估计和自助法分布：英国股票收益型基金

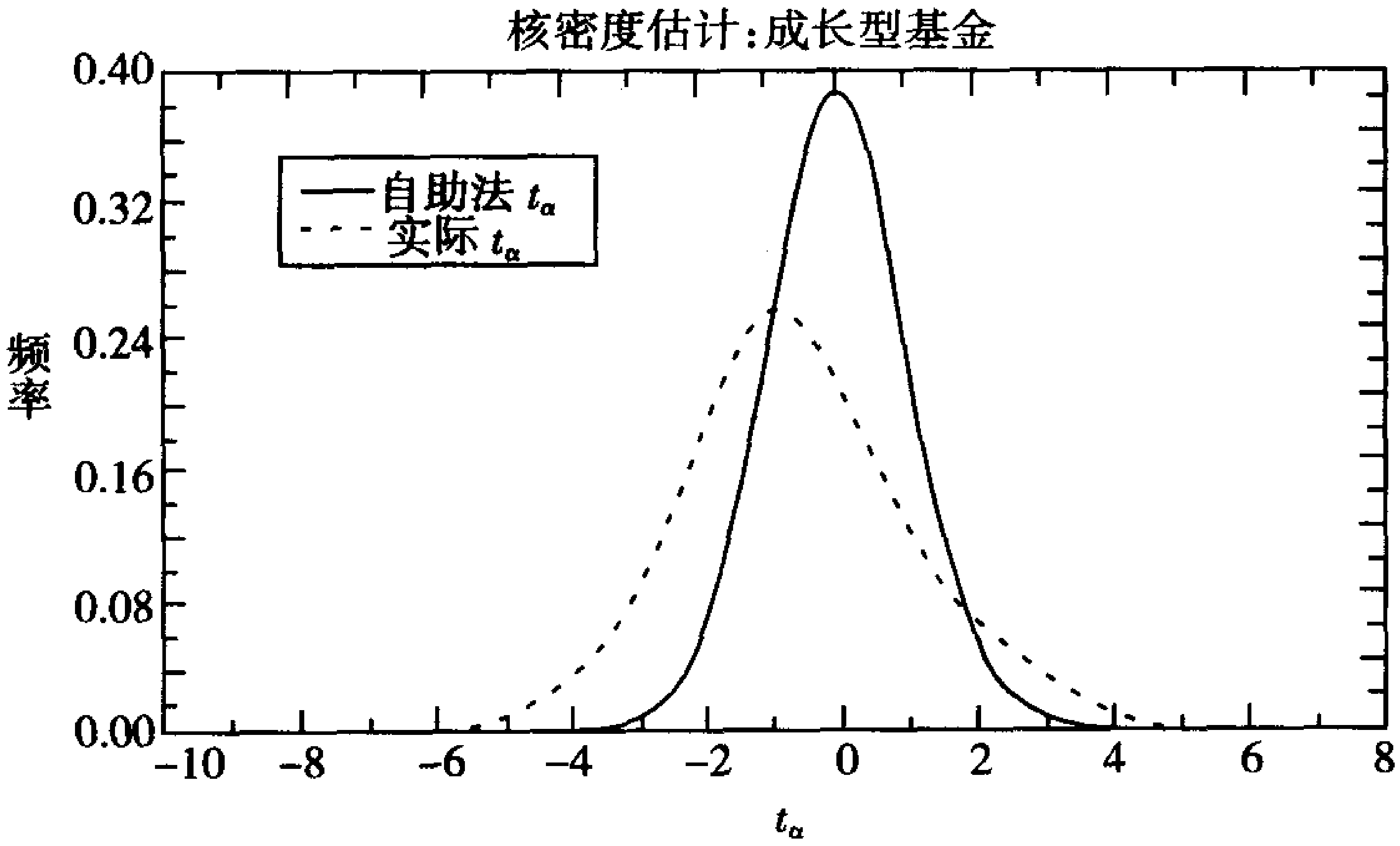


图6 实际 t_α 的核密度估计和自助法分布：英国股票成长型基金

基金的历史长度

将自助法分析限制到观测值数目最少的基金上或许是很有意义的，这种做法可以将运行时间很短的那些基金排除在外，对这些基金而言，所得到的 α 绩效估计值具有较大的样本差异，这使得样本估计结果变得不可靠。估计结果不可靠的基金在实际分布和自助法分布的极端尾部中不成比例分布。选择而不是构造 t_α 的自助法 p 值可以减轻这一问题的影响。即使这样，为了避免修正后的绩效分布和未修正的绩效分布出现这样的问题，对观测值的最小数目进行限制是有意义的。这一节的所有自助法 p 值都是用样本观测值最小数目为60的基金来进行计算的。然而，这一方法的可能缺陷是，它可能引入幸存偏好问题，因为分析被限制在那些因为技能（或运气）而至少存活五年的基金上。

因为存在这种偏差来源，所以考察到目前为止的所有结论对最小观测值数目的敏感性是必要的。为了对此进行考察，这里的研究使用自助法程序，对基金样本容量的各种可能取值重新估计 p 值。在所有情形中使用的都是无条件三因子模型，结论与包括所有风格基金的完整样本有关。上面所陈述的结论对样本容量不敏感（即幸存偏好问题不重要）。特别地，各组分布左尾的绩效都拒绝用偶然因素来解释糟糕的基金绩效，但对右尾而言，好运气常常是导致“良好”绩效的原因。

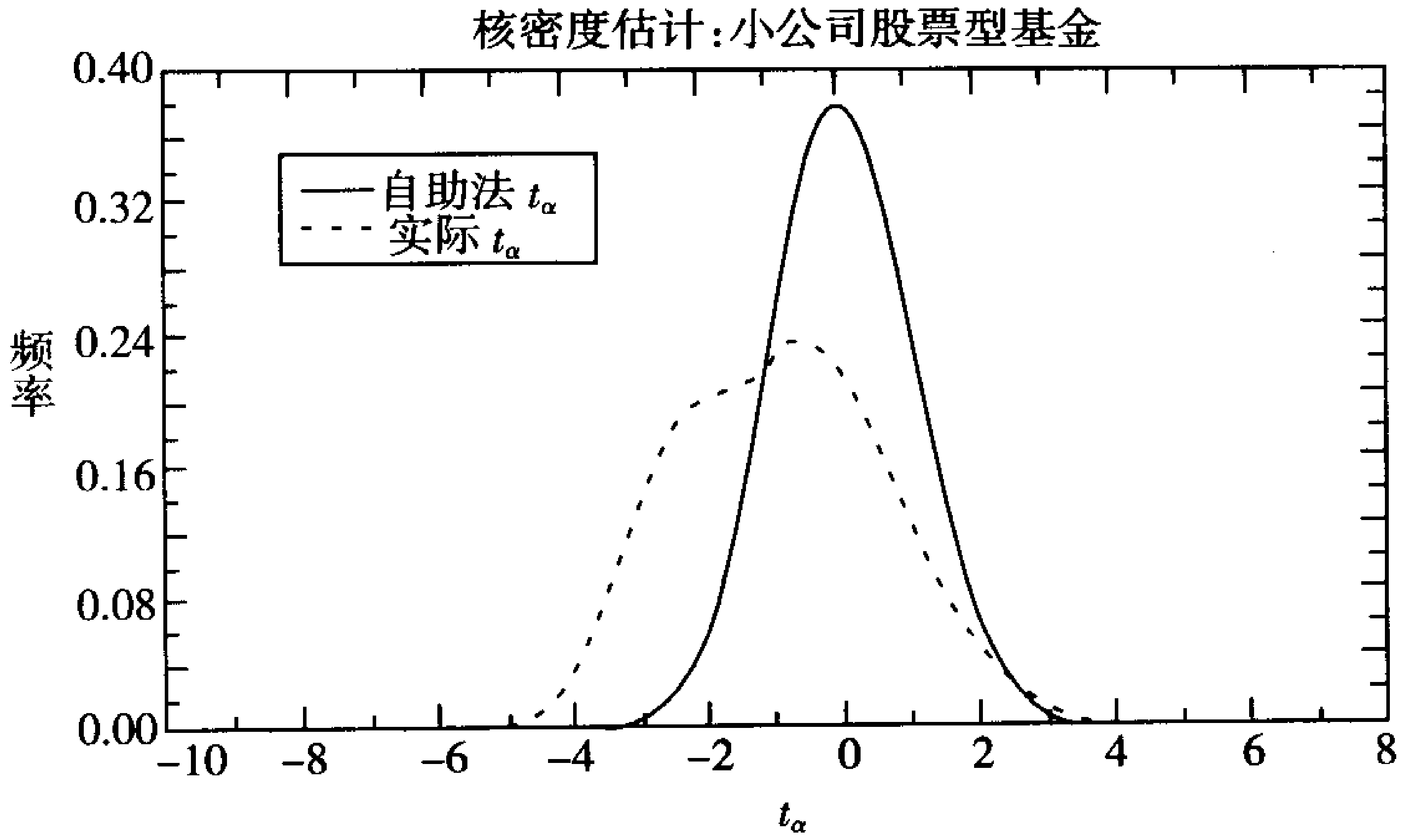


图 7　实际 t_α 的核密度估计和自助法分布：英国普通股票型基金

残差时间序列的依赖性

为了选择与可能阶数的序列自相关性相对应的残差样本容量，我们可以对自助法程序进行修正。这里的研究表明，大部分基金都无法拒绝一阶序列相关性的零假设。我们对可能的残差样本容量重复使用自助法模拟，但 p 值基本上没有什么变化。因此，自助法主要是针对分布尾部的非正态性而不是针对残差中的序列相关性而言的。

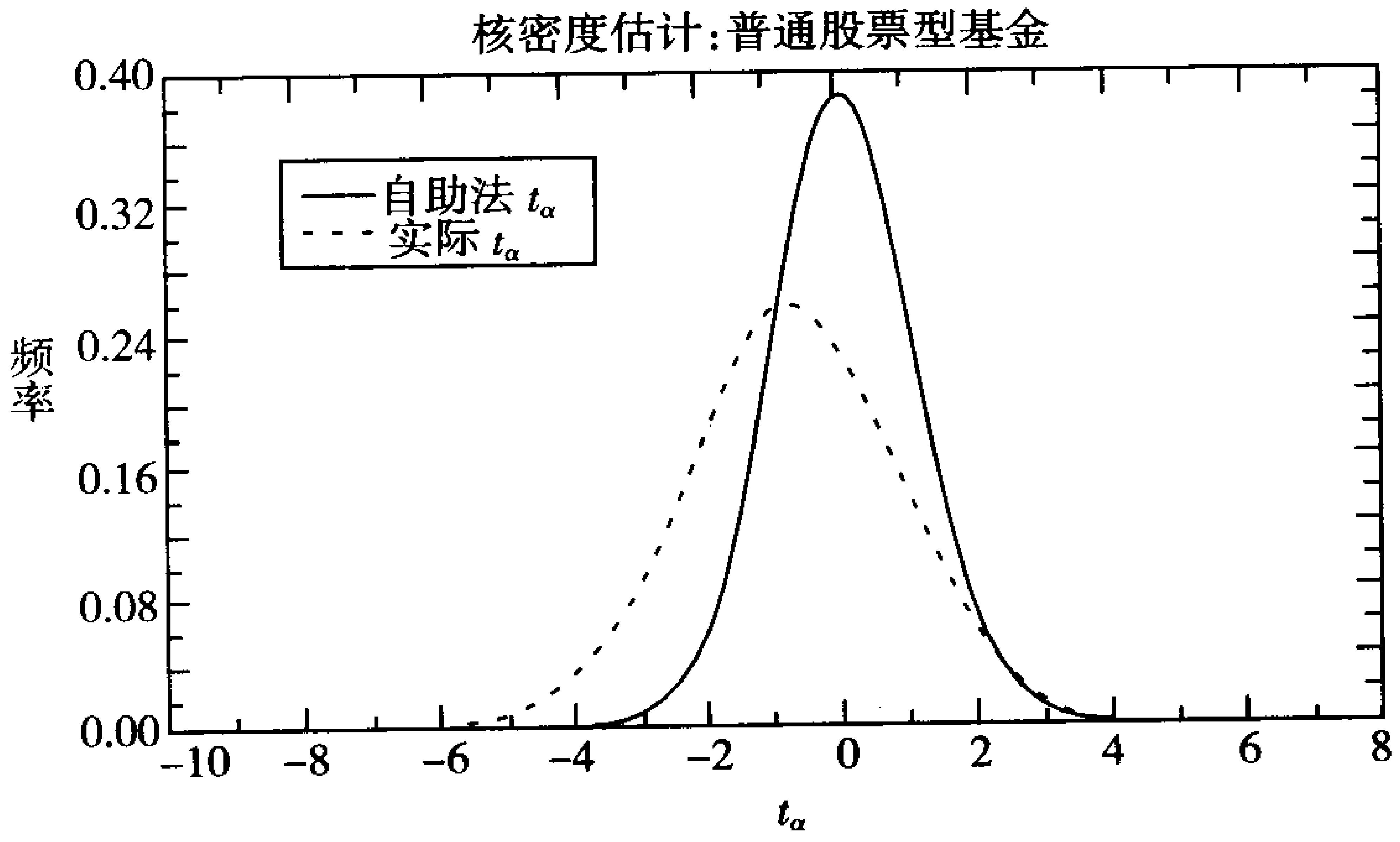


图 8　实际 t_α 的核密度估计和自助法分布：英国小公司股票型基金

总的来说，对英国共同基金研究的初步结果表明，在英国绩效最好的共同基金中，几乎没有哪只基金是因为选股能力而获得良好的绩效，而基金绩效较差不是因为运气不好而是因为技能太低。当然，大部分基金的表现都很平庸，这里的研究没有强调绩效持续性问题。所报告的结果都是针对无条件三因子 Fama - French 模型而言的，对各种绩效模型（如条件模型和择时交易问题）的敏感性分析还需要更为深入的考察。

9.4　小结

* 事件研究使用各种市场模型来检验事件宣告（例如敌意并购）对股票异常收益的影响，CAR 常被用作衡量市场对这家公司（股票）价值信息的看法。

* 多因子方法常用于对共同（养老）基金绩效的研究，探寻有可能影响共同基金组合收益的风险因素，而在对异常收益进行度量时，Jensenα 是一个常用方法，这一方法也用于对绩效持续性的研究。

* 对 α 值进行标准统计检验表明在分布的两端存在绩效特别好和特别差的基金，但大部分还是很平庸。所有良好绩效都只能持续一两年，但糟糕绩效却保持得更久一些。

* 使用自助法分析排除了运气因素后，在对美国和英国的研究中都发现分布的右尾确实有一些绩效显著的基金（在美国的研究中表现得更为显著），然而两国的研究也同时揭示了左尾中的一些基金绩效差的原因，即选股能力确实差劲。然而，这一方法相对来说还比较新颖，因此，这些开创性结果尽管很有意思，但在运用时仍需谨慎。

第10章

价值评估模型与资产收益

目的

* 在预期股票收益为常数时证明股票价格等于未来红利的贴现值（DPV），贴现率为常数。这就是理性定价公式（RVF）。

* RVF 的一种特殊情形（红利增长率和贴现率都为常数）就是 Gordon 增长模型。

* 如果均衡预期收益随着时间发生改变，那么股票价格 RVF 中的贴现率也具有时变性。

* 说明如何对具有时变贴现率和时变红利增长率的 RVF 进行线性化处理，分析股票价格波动率与收益可预测性之间的联系。

这一章关注投资者用来决定某只股票基础价值或公平价值（V_t）的模型，决策的“基础”是红利和未来必要收益。本章的核心思想是，股票收益与股票价格之间有着解不开的联系。事实上，各种均衡预期收益模型对股票基础价值（从而股票价格）的表达形式也不同。

10.1　理性定价公式

预期收益的定义是：

$$E_tR_{t+1} = \frac{E_tV_{t+1} - V_t + E_tD_{t+1}}{V_t} \tag{1}$$

V_t 是股票在第 t 期末的价值，D_{t+1}是 t 到 $t+1$ 期之间的红利支付，E_t 是基于 t 时刻或之前信息集合 Ω_t 的期望算子，

$$E\ (D_{t+1} \mid \Omega_t) \equiv E_tD_{t+1}$$

我们假设，只要股票收益的预期收益为常数（$=k$），投资者就愿意持有这只股票。就目前而言，我们可以将“必要收益”k 视为足以补偿投资者承担股票内在风险的收益，但我们不对 k 进行“模型”设定，

$$E_tR_{t+1} = k \qquad (k>0) \tag{2}$$

$R_{t+1} - k$ 的随机行为使得平均意义下的异常收益不存在，超额收益是一种“公平赌博”，

$$E_t\ (R_{t+1} - k \mid \Omega_t) = 0 \tag{3}$$

根据方程（1）和（2）就可以得到决定“价值”运动的欧拉方程：

$$V_t = \delta E_t\ (V_{t+1} + D_{t+1}) \tag{4}$$

贴现因子为 $\delta = 1/\ (1+k)$，$0<\delta<1$。将方程（4）向前滚动一期就得到

$$V_{t+1} = \delta E_{t+1}\ (V_{t+2} + D_{t+1}) \tag{5}$$

假设信息的获取时间只能是 t，对（5）求数学期望就得到

$$E_t V_{t+1} = \delta E_t\ (V_{t+2} + D_{t+2}) \tag{6}$$

在导出方程（6）的过程中，我们使用了迭代期望法则

$$E_t\ (E_{t+1} V_{t+2}) = E_t V_{t+2} \tag{7}$$

方程（7）的左边可以解释为在今天对明天（$t+1$）关于 V_{t+2} 的预测值进行预测，它等于今天对 V_{t+2} 的预测，因为在理性预期假设下，你不知道你的未来预期会发生怎样的变化。因为方程（6）对所有时期都成立，所以

$$E_t V_{t+2} = \delta E_t\ (V_{t+3} + D_{t+3}) \tag{8}$$

其余类似。将（6）代入（4）得到解的下一个部分

$$V_t = \delta\ [\delta E_t\ (V_{t+2} + D_{t+2})] + \delta\ (E_t D_{t+1})$$

连续迭代后就得到

$$V_t = E_t\ [\delta D_{t+1} + \delta^2 D_{t+2} + \delta^3 D_{t+3} + \cdots + \delta^N\ (D_{t+N} + V_{t+N})] \tag{9}$$

现在令 $N \to \infty$，因此 $\delta^N \to 0$。如果 D 的预期增长不发散且使 $E_t V_{t+N}$ 为有限数，那么

$$\underbrace{\lim}_{n\to\infty} E_t \delta^N\ [D_{t+N} + V_{t+n}] \to 0 \tag{10}$$

方程（10）就是众所周知的终端条件（terminal condition）或横截性条件（transversality condition），它将理性投机泡沫（rational speculative bubbles，见 17 章）排除在外。于是方程（9）变为：

$$V_t = E_t \sum_{i=1}^{\infty} \delta^i D_{t+i} \tag{11}$$

我们在如下假设下导出了方程（11）：预期收益为常数，迭代期望法则对所有投资者都成立，横截性条件成立。

股票的基础价值 V_t 是预期的未来红利贴现值，如果 $P_t \neq V_t$，那么市场上就存在未被利用的获利机会。例如，如果 $P_t < V_t$，那么投资者可以买入股票，因为他们的预期是，当股票价格 P_t 向"正确"价值靠近时能获得资本收益。当 $P_t < V_t$ 时投资者就可以购买股票，需求增加会导致股票价格上升，价格会迅速地向基础价值方向运动。现在假设：

* 边界上的投资者具有同质预期，更为准确地说，他们对基础价值概率分布的主观看法是"真实"内在分布的反映。

* 无风险套利瞬间就可以完成，所有投资者认为实际市场价格 P_t 等于基础价值 V_t。

于是，我们就得到了贴现率为常数时股票价格的 RVF：

$$P_t = E_t [\sum_{i=1}^{\infty} \delta^i D_{t+i}] \tag{12}$$

在上面的分析中，我们没有对实际变量和名义变量进行区分，事实上，数学公式在两种情形中都成立。因此，无论变量是名义变量还是用总商品价格指数进行平减后的变量，方程（12）都成立。然而，不管是根据直观推理还是根据实证研究结果，预期实际收益都比预期名义收益（至少是由通货膨胀所引起的）更加平稳。因此，通常用实际变量来表示 RVF。

有限或无限时期？

如果投资者考虑的是一个有限时期，那么他们关注的是不久就可以获得的股票价格，这是否会改变我们对股票基础价值决定因素的判断呢？考虑一个简单的单期情形：

$$P_t = \delta E_t D_{t+1} + \delta E_t P_{t+1}$$

今天的价格取决于对 $t+1$ 时刻的预期价格。但是，如果投资者要在 $t+1$ 时刻卖出股票，那么他该如何判断 $E_t P_{t+1}$ 的值呢？如果投资者是理性的，那么他将采取与价格 P_t 相同的决定方式来进行判断，即

$$E_t P_{t+1} = \delta E_t D_{t+2} + \delta E_t P_{t+2}$$

重复前面的推导过程。单期中的每一个投资者都相信 P_{t+j} 由上述欧拉方程来决定，因此今天的价格等于未来所有时期中红利的 DPV。因此，即使有些投资者只着眼于有限时期，在决定股

票基础价值方面，他们也与无限期限中的投资者一样。一种替代观点是，如果投资者都是短期持有股票，那么价格就不能反映基础价值，这就是众所周知的短视主义（short-termism）。

本章余下部分对均衡收益为常数的假设进行放松，我们发现 RVF 的形式仍然有效。与此相对照的是，在第 18 和 19 章中，非理性投资者或噪声交易者会对收益产生影响。在那些模型中，股票价格不再等于基础价值。

10.2 理性定价公式的特殊情形

A. 预期红利为常数

假设未来（实际）红利的最佳预测就是红利的当前值，于是

$$D_{t+1}=D_t+w_{t+1}$$

这里的 w_{t+1} 为白噪声，红利遵循随机游走过程。在理性预期假设下，$E_t(w_{t+j}\mid\Omega_t)=0$，$E_tD_{t+j}=D_t$（$j\geqslant 1$）。因此，预期的红利增长率为零，理性定价公式（12）变成

$$P_t=\delta(1+\delta+\delta^2+\cdots)D_t=\delta(1-\delta)^{-1}D_t=(1/k)D_t \tag{13}$$

方程（13）意味着红利价格比（D/P）或红利收益比为常数，且等于必要（真实）收益 k。模型（13）也意味着股票价格变化的百分比等于红利变化的百分比。只有在获得红利（即“基础变量”）方面的新信息后股票价格才会发生改变，模型意味着股票价格波动率完全取决于红利波动率。尽管投资者对未来各期红利的最佳预测是 D_t，随机因素 w_{t+1}（即消息或不确定性）可能导致未来红利实际值偏离投资者的最佳预测。因此，价格的条件方差取决于这些基础变量方面的消息的方差：

$$P_{t+1}-P_t=(1/k)\Delta D_{t+1}=(1/k)w_{t+1} \tag{14}$$

$$\mathrm{Var}(P_{t+1}-P_t\mid\Omega_t)=(\sigma_w/k)^2 \tag{15}$$

在本书后面部分中（第 12 章），我们将考察更为一般的情形，即股票价格波动率不仅取决于红利波动率，还取决于贴现率波动率。事实上，我们还将试图回答股票价格波动率主要取决于红利波动率还是贴现率波动率。

B. Gordon 增长模型

（实际）红利以常数 g 进行增长的时间序列模型就是自回归模型 AR（1）：

$$D_{t+1}=(1+g)D_t+w_{t+1} \tag{16}$$

这里的 w_{t+1} 是白噪声。由方程（16）很容易看出预期红利增长率等于 g

$$(E_tD_{t+1}-D_t)/D_t=g \tag{17}$$

注意，如果红利的对数服从随机游走且漂移系数为 g^*，那么这也可以得到红利的常数预期增长率（即 $E_t(\ln D_{t+1})=g^*+\ln D_t$）。将（16）反复代入就可以得到未来红利的最优预测值

$$E_tD_{t+j}=(1+g)^jD_t \tag{18}$$

将（18）中的未来红利预测值代入理性定价公式就得到

$$P_t=\sum_{i=1}^{\infty}\delta^i(1+g)^iD_t \tag{19}$$

经过简单的代数运算就得到 Gordon 增长模型

$$P_t=\frac{(1+g)}{(k-g)}D_t\qquad((k-g)>0) \tag{20}$$

例如，如果 $g=0.03$，$k=0.08$，那么价格红利比为 20.6。如果投资者突然改变对 g 或 k 的预期，那么股票价格将发生很大的变化。例如，如果 g 降到 0.02，那么新的价格红利比为 17，这意味着价格会“立即”下降 17.4%。如果我们假定 g 和 k 仍为常数，那么模型在短期内无

法“拟合事实”，因为在短期内（如五年以内）价格波动率比红利波动率大得多，尽管在更长的时期内（如25年或更长）价格和红利倾向于一起运动（即协整关系），这与模型预测的结果相同（即 $\ln P_t=\alpha+\ln D_t$，其中 $\alpha=\ln[(1+g)/(k-g)]$）。

10.3　时变预期收益

为了使投资者愿意持有某只特殊的股票，我们假设投资者在未来各个时期都要求一个不同的预期收益。（后面将对这种情形发生的原因进行考察。）因此，我们有

$$E_tR_{t+1}=k_{t+1} \tag{21}$$

其中，k 是时间下标，表示时变性。重复上述步骤并逐步向前替代就得到

$$P_t=E_t\ [\delta_{t+1}D\delta_{t+1}+\delta_{t+1}\delta_{t+2}D_{t+2}+\cdots+\delta_{t+1}\cdots\delta_{t+N-1}\delta_{t+N}\ (D_{t+N}+P_{t+N})] \tag{22}$$

上式可以写成一种紧凑的形式（假设横截性条件成立）

$$P_t\ =\ E_t\Big[\sum_{j=1}^{\infty}\Big[\prod_{i=1}^{j}\delta_{t+i}\Big]D_{t+j}\Big]\ \equiv E_t\sum_{j=1}^{\infty}\delta_{t,t+j}D_{t+j} \tag{23}$$

其中，$\delta_{t+i}=1/(1+k_{t+i})$，$\delta_{t,t+j}=\delta_{t+1}\cdots\delta_{t+j}$。因此，当前的股票价格取决于未来贴现率和红利乘积的总和的预期值。注意，对所有时期而言，$0<\delta_{t+j}<1$，越远的预期红利对当前股票价格的影响越小。然而，可能的情形是，今天的某个事件宣告（如公司并购）对（比方说）五年后的红利会产生重大影响。在这种情形中，尽管贴现率较大一些，事件公告也会对当前的股票价格产生重大影响。注意，在一个信息充分（有效）的市场上，尽管五年之内没有任何红利支付，人们期望股票价格立即对事件宣告产生全面反应。与此相对照的是，如果市场无效率（如存在噪声交易者），那么不仅当期价格会上涨，以后各期的价格也会上涨。股票价格对事件宣告的反应就是众所周知的事件研究。

到目前为止我们还无法对（23）“进行计算”，因为它涉及不可观测的预期项。我们无法计算基础价值（即（23）的右边），因此也就无法观察基础价值是否与当前价格 P_t 一致。就投资者对红利和贴现率的预测而言，我们需要补充易于处理的假设。相对直接的方法是发展红利预测方程，例如，就年度数据而言，AR（1）和 AR（2）模型对红利数据的拟合效果非常好。困难在于均衡收益的确定。我们将在后面的章节中对两期和多期均衡收益模型进行考察，导出均衡收益对股票价格和收益的含义。

就目前而言，我们将 CAPM 模型作为考察均衡收益的一种特殊方法，方法之所以特殊是因为 CAPM 模型的基础是单期均值方差优化问题，而我们是在多期 DPV 公式中使用 CAPM 模型。我们的用法在逻辑上是不一致的。但是，在应用工作中这种特殊方法是众所周知的。

市场投资组合

根据单期 CAPM 模型，所有投资者在均衡条件下都持有市场投资组合（即每个投资者对风险资产的持有比例与市场投资组合的比例相同）。Merton（1973）在跨期框架中对这一思想进行了发展，证明了如下结论：市场投资组合相对于无风险利率的超额收益与市场投资组合收益方差的预期值成比例

$$E_tR_{m,t+1}-r_t=\lambda\ (E_t\sigma^2_{m,t+1}) \tag{24}$$

预期收益定义为无风险收益加上风险溢价 rp_t：

$$E_tR_{m,t+1}=r_t+rp_t \tag{25}$$

$$rp_t=\lambda E_t\sigma^2_{m,t+1}$$

根据 CAPM 模型，对（21）和（25）进行比较就可以看出市场投资组合的必要收益为

$$k_t=r_t+\lambda\ (E_t\sigma^2_{m,t+1}) \tag{26}$$

均衡必要收益与无风险收益 r_t 和市场投资组合的（不可分散的）风险正相关，市场投资组合

的风险用条件方差 $E_t\sigma^2_{m,t+1}$ 来进行度量。如果

- $*$ 投资者不认为市场存在风险（即 $E_t\sigma^2_{m,t+1}=0$），或
- $*$ 投资者属于风险中性类型（即 $\lambda=0$），

那么投资者将无风险利率 r_t 用作贴现因子是合适的。注意，为了用 RVF 来确定资产的价格，投资者必须确定 k_t 的值，需要对无风险利率和风险溢价进行预测。

个体资产的收益

现在考虑单个证券或资产组合的价格，资产组合中的资产是市场投资组合中资产的一个部分（如所有工业股票或所有银行股票）。CAPM 模型意味着，为了使投资者持有分散化投资组合的一个部分，投资组合 i 的预期收益应为

$$E_tR_{i,t+1}=r_t+\beta_{it}\ (E_tR_{m,t+1}-r_t) \tag{27}$$

$$\beta_{it}=E_t\ (\sigma_{im}/\sigma^2_m)_{t+1}$$

将（24）代入后就得到

$$E_tR_{i,t+1}=r_t+\lambda E_t\ (\sigma_{im})_{t+1} \tag{28}$$

这里的 σ_{im} 是资产 i 收益和市场投资组合收益之间的协方差。将（21）与（28）进行比较就可以看出，资产 i 的均衡必要收益为

$$k_{t+1}=r_t+\lambda E_t\ (\sigma_{im})_{t+1}$$

协方差项可能具有时变性，因此在单个证券（或证券组合）的 RVF 中，未来贴现因子 δ_{t+j} 也可能具有时变性。在 CAPM 模型中，市场只对投资者所承担的不可分散（系统）风险给予回报。投资者自愿持有宽泛投资组合中部分资产的必要（名义）收益 k_t 等于无风险利率加上风险回报或风险溢价 rp_t，风险溢价取决于（条件）协方差。

理性定价公式的线性化

正如在 12 章将要看到的那样，对具有时变贴现率的 RVF 进行线性化是可能的，在股票价格运动的分析中线性化后的公式非常有用。为了“对此进行处理”，考虑用方程（20）表示的 Gordon 增长模型，模型意味着目前高的价格红利比要么归因于高的预期红利增长率，要么归因于低的未来收益（贴现率）。现在考虑对 RVF 进行如下线性化处理（假设理性泡沫不存在）

$$p_t-d_t=\text{常数}+E_t\ \{\sum_{j=1}^{\infty}\rho^{j-1}\ (\Delta d_{t+j}-h_{t+j})\}\ +\lim_{j\to\infty}\rho^j\ (p_{t+j}-d_{t+j}) \tag{29}$$

其中 ρ 是线性化过程中的常数，它的取值约为 0.95，Δd_{t+j} 为红利对数值的变化，是单期收益的对数值序列。方程（29）是 Gordon 增长模型的动态版本，红利增长率和预期收益都不是常数，在不同时期中取不同的值。这与 Gordon 增长模型是一致的，因为价格红利比与红利增长率正相关，与未来预期收益负相关。假定红利价格比随着时间而发生变化，根据（29），它可以对未来红利增长率或未来收益进行预测。换句话说，如果 RVF 在真实世界中有效，那么高价格红利比之后必然是（平均意义下）高红利增长率或低收益 h_{t+j}（即贴现因子）。

上述方程隐含着许多问题，我们在后面的章节中进行讨论。下面两章将要讨论 Shiller 的工作：实际价格（或价格红利比）的波动率不能简单地用理性投资者对未来红利增长率预期的变化来进行解释，因为从数据来看后者的变化不是很大。因为利率发生了变化，所以预期股票收益 E_th_{t+j} 会发生变化，这样的假设无助于矛盾化解，因为这个假设意味着高的价格红利比预示着更低的利率（平均而言），但数据显示的结果却不是这样。因此，RVF 的接受在很大程度上依赖于风险规避程度的变化或股票市场风险的变化。实证研究表明，股票市场的波动率与预期收益（见（28））之间不存在很强的联系（如 French、Schwert 和 Stambaugh（1987）），我们在第 29 章中使用所谓的 GARCH－M（GARCH－in－Mean）模型来对相关证据进行考察。在第 13 章中，我们在随机折现因子的框架下讨论股票市场风险的变化，风险规避程度的变化是第 18 和 19 章中“行为金融学”模型讨论的主题。（29）的最后一项也可以解释为价格红利比的

运动，这是第17章理性泡沫所讨论的主题（参见 Brunnermeier（2001））。

当然，我们也可以认为投资者并不知道真实的经济结构（可以将其看作一组白噪声），投资者需要花费时间来了解经济环境。这一思路也许可以解决波动率之谜，因为短期价格的随机上涨会在未来产生影响，导致未来价格发生更大波动，波动幅度比真正理性的投资者所预测的还要大。信仰改变和“不正确的”预期形成这一思想贯穿全书，尤其是第17章。

波动率与长期可预测性

如果正如数据所反映的那样，价格红利比不能预测未来红利，那么它就一定可以预测未来收益。因此，理性定价公式（29）与使用红利价格比来预测长期收益所得到的回归结果是一致的，$\sum_{j=1}^{\infty}\rho^{j-1}h_{t+j}$更像是长期收益（因为$\rho$接近于1）。因此，从数据来看，高的价格红利比似乎不能反映更高的红利增长率，而对更低的预期和预期风险溢价有所反映。

线性化 RVF 的另一种表示是

$$h_t - E_{t-1}h_t = (E_t - E_{t-1})\left\{\sum_{j=0}^{\infty}\rho^j(\Delta d_{t+j} - \Delta h_{t+j+1})\right\} \tag{30}$$

$(E_t - E_{t-1})$表示预期的变化或“意外”事件。因此，单期收益的意外变化要么是因为对未来红利的预期进行了修正，要么是因为对未来收益的预期进行了修正（或两者兼有）。Campbell（1991）发现。对未来收益（贴现率）预期的修正是导致当前收益意外变化的主要原因，对收益的影响而言，红利处于一个相对次要的地位。方程（30）也说明，相反的直觉与如下结果在逻辑上是一致的：如果红利方面的信息没有变化，那么预期未来收益的上升意味着当前收益会出现预料不到的下降。这是因为，如果预期未来贴现率增加但红利不发生改变，那么今天的价格就必然下降。

预期收益与价格波动

在实证研究中，股票价格较大幅度的波动与预期收益相对较小的可预测变化之间是否存在一致性呢？在预期收益具有持续性的假设下，答案是肯定的。由 Gordon 增长模型可以很容易地发现，必要收益k持久的微小变化会导致股票价格的大幅变化。现在，假定预期收益服从 AR（1）过程，系数ϕ接近于1

$$h_{t+1} = \phi h_t + \varepsilon_{t+1} \tag{31}$$

注意，（31）意味着$E_t h_{t+2} = \phi E_t h_{t+1}$，从而预期收益具有持久性。将$E_t h_{t+j} = \phi^j h_t$代入（29）（忽略红利增长项）就得到

$$(p_t - d_t) = \frac{-1}{(1-\rho\phi)}h_t \tag{32}$$

合理的参数取值为$\rho = 0.95$，$\phi = 0.9$。预期收益的“微小”波动$\sigma(E_t h_{t+1}) = 1\%$会导致股价$\sigma(p_t - d_t) = 6.9\%$的大幅变动。我们将在第12章中对这个问题进行深入讨论。

10.4　小结

* 在瞬时套利的假定下，当预期收益模型给定时，我们可以证明股票价格等于红利与贴现因子乘积的贴现预期值。这就是理性定价公式（RVF）。如果均衡预期收益为常数，那么理性定价公式中的贴现率也为常数。

* Gordon 增长模型假定贴现率和红利增长率都为常数，即红利和贴现率都具有持续性。股票价格对贴现率或红利增长率的微小变化非常敏感。如果贴现率和红利增长率在所有时期中都为常数，那么（对数）价格和（对数）红利之间是一一对应关系。

* 如果资产的均衡预期收益由标准 CAPM 模型给定，那么理性定价公式中的贴现因子可

能具有时变性，贴现因子依赖于无风险利率和（市场收益的）方差项或协方差项（单个资产或市场投资组合中资产的某个子集）。

* 假设红利增长与未来贴现率是时变的，红利增长不可预测（如（对数）红利服从随机游走模型），即使预期收益勉强可以预测且具有持续性，股票价格也可能有大的波动率。

第 11 章

股票价格的波动率

目的

- 说明如何用 Shiller 的方差边界检验方法来对股票价格的理性定价公式进行评价，这就导致了波动率之谜（volatility puzzle）。
- 分析小样本偏误的统计问题和方差边界检验中的非平稳性。
- 说明“比索问题”如何改变我们对方差边界检验的解释。
- 说明方差边界检验与 RVF 各种回归检验之间的联系。

在这一章里，我们讨论“基础变量”（如红利或未来收益预测值的变化等）是否能解释所观察到的股票价格波动率。如果理性基础变量模型（rational fundamental model）不能解释所观察到的波动率，那么我们就得到了一种异像，这就是众所周知的波动率之谜（volatility puzzle, Campbell（2000））。在后面的章节里我们将看到与股票收益有关的其他典型事实，这些典型事实无法用传统模型来进行合理解释，最为著名的就是股权溢价之谜（equity premium puzzle）和无风险利率之谜（risk－free rate puzzle）。

波动率检验是在均衡预期收益的某个特殊假设下对股票价格的理性定价公（RVF）直接进行考察。最简单的假设是单期收益为常数。

$$P_t = \sum_{i=1}^{\infty} \delta^i E_t D_{t+1} \tag{1}$$

如果能获得预期红利的可靠度量，那么我们就能计算出方程（1）的右边，于是对 RVF 的检验可以用来考察 $\mathrm{Var}(\sum \delta^i E_t D_{t+i})$ 与 $\mathrm{Var}(P_t)$ 是否相等。Shiller（1981）的学术论文避免了预期红利数据的使用问题。他注意到，在理性预期（RE）下，实际红利与预期红利之间的差异只是一个随机（预测）误差，因此，实际股票价格 P_t 与完全预预见价格（the perfect foresight price）P_t^*（$= \sum \delta^i D_{t+i}$）之间的差异也只是一个随机误差。注意，P_t^* 可以使用实际的红利值。Shiller 认为 RVF 意味着 $\mathrm{Var}(P_t) \leqslant \mathrm{Var}(P_t^*)$。与此同时，另外两位金融经济学家 LeRoy 和 Porter（1981）也进行了类似检验。然而，Shiller 对美国股票价格进行研究后发现，方差不等式总体来看不成立，LeRoy 和 Porter 发现方差不等式（在统计意义上）被拒绝的概率却很小。在这两篇文献的基础上，对基础方法的各种变体进行研究的文献大量涌现。有些文章强调小样本偏差，最近的一些研究在红利序列不是稳定过程的假设下考察波动率检验的稳健性。

许多评论家经常陈述的观点是股票市场有着过度波动性，即价格每日或每周的大幅波动并没有反映基本面上“消息”的变化。如果情形确实如此，那么这就否定了有效市场假说（EMH）。当然，如果要证明股票价格产生了过度波动，我们需要以理性行为为基础的模型，这个模型是我们进行比较的基准。当然，常识告诉我们股票价格具有一定程度的波动性，因为

公司方面的消息或“新信息”随时可能出现。然而，我在这里要强调的问题不是股票价格是否具有波动性，而是股票价格是否具有过度波动性。

现在，我们在方差边界的基础上来对 RVF 的规范检验方法进行考察。从宽泛意义上讲，方差边界检验考察的是股票价格的方差是否与 RVF 所给出的基础变量（即红利与贴现率）的变化一致。这种检验方法可以分为两种大的类型：“不以模型为基础”的检验和“以模型为基础”的检验。就前者而言，我们不用假设基础变量服从某个特殊统计模型。然而，正如我们即将看到的那样，这意味着我们只能获得相关检验统计量的点估计，无法得到估计值的置信区间，因此无法进行规范的假设检验。我们能做的就是试图确保合意统计量（基于样本）的估计值是总体相应值的无偏估计。对早期方差比检验的批评强调的就是有限样本中的偏差问题（Flavin（1983））。

“以模型为基础的检验”假设红利服从某个特殊的随机过程，给出了检验统计量合适的置信区间，这使得我们能用蒙特卡罗模拟来对小样本性质进行考察。然而，对以模型为基础的检验而言，RVF 的检验前提是红利和贴现率的相关模型是正确的模型。因此，我们就面临联合零假设问题。

对股票价格各种检验进行解释的一个关键因素是红利过程是否是平稳过程。如果红利过程非平稳，那么通常的分布假设就不适用，对方差边界检验结果的解释也存在很大问题。这一领域的近期工作致力于对非平稳数据进行检验的程序。

11.1 Shiller 波动率检验

在常数（实际）收益假设下我们有

$$P_t = \sum_{i=1}^{n} \delta^i E_t D_{t+i} + \delta^n E_t P_{t+n} \tag{2}$$

P_{t+n}表示在 $t+n$ 时刻的“最终价格”，所有投资者对未来持有相同的看法。我们在 t 时刻不知道投资者对未来红利的预测值。然而，Shiller（1981）对这一问题提出了一种简单而又巧妙的解决方法。假设我们有实际红利的历史数据，比方说 1900 年以来的所有数据；也有当前的实际价格 P_{t+n}，比方说 2004 年的实际价格。假设 δ 的值已知，比方说 0.95，就年度数据而言，这意味着必要收益为 $k=5.2\%$（$\delta=1/(1+k)$）。于是，如果投资者有 1900 年以来红利的准确预测值，那么根据方程（2）就可以计算出 1900 年的股票价格。我们将这个价格 P_t^* 称为 1900 年的完美预见价格。将年份往前推移一年并重复上述程序，使用下面的公式就可以获得 1900 年之前各个 P_t^* 的时间序列数据。

$$P_t^* = \sum_{i=1}^{n} \delta^i D_{t+i} + \delta^n P_{t+n} \tag{3}$$

在计算 1900 年的 P_t^* 时，最终价格 P_{t+n}的影响最小，因为 n 越大，δ^n 就越小。当我们接近终点时，比如说在计算 2004 年的 P_t^* 时，$\delta^n P_{t+n}$有着更大的权重。一种选择是对样本进行截断，比方说在 10 年前的位置进行截断。另外一种选择是假设终点时刻的实际价格“接近于”预期值 $E_t P_{t+N}$。实证研究经常采用后面一种方法。比较 P_t 和 P_t^* 就可以看出两者之差等于红利预测误差 $\omega_{t+i} - E_t D_{t+i}$之和，权重为贴现因子 δ^i。

由（2）、（3）及 ω_{t+i}的定义可知，

$$P_t^* = P_t + \sum_{i=1}^{n} \delta^i \omega_{t+i} + \delta^n (P_{t+n} - E_t P_{t+n})$$

假设最后一项在 $n \to \infty$ 时接近于零，注意在理性预期假设下 Cov $(P_t, \omega_{t+i}) = 0$，于是 RVF 意味着 Var $(P_t^*) \geqslant$ Var (P_t)。

当 $x_t = P_t$ 或 P_t^* 时，样本方差为 Var $(x) = \sum_{t=1}^{n}(x_t - \bar{x})^2/(n-1)$，$\bar{x}$ 表示样本均值，n

是样本观测值的数目。投资者在 1900 年不知道未来红利为多少，因此，实际股票价格与完全预见股票价格不同。从事后来看，我们知道投资者的预测误差为 $\eta_t = P_t^* - P_t$，因此

$$P_t^* = P_t + \eta_t \tag{4}$$

其中，$\eta_t = \sum_{i=1}^{n} \omega_{t+i}$。在理性预期下，$\eta_t$ 独立于 t 时刻的所有信息，特别地，η_t 独立于 t 时刻的股票价格。由方程（4）可以得到

$$\text{Var}\ (P_t^*) = \text{Var}\ (P_t) + \text{Var}\ (\eta_t) + 2\text{Cov}\ (\eta_t,\ P_t) \tag{5}$$

信息有效（正交性）意味着 Cov（η_t，P_t）等于零，因此

$$\text{Var}\ (P_t^*) = \text{Var}\ (P_t) + \text{Var}\ (\eta_t) \tag{6}$$

因为 Var（η_t）≥0，所以

$$\text{Var}\ (P_t^*) \geqslant \text{Var}\ (P_t) \tag{7a}$$

或

$$\text{VR} = \text{Var}\ (P_t^*) / \text{Var}\ (P_t) \geqslant 1$$

$$\text{SDR} = \sigma\ (P_t^*) / \sigma\ (P_t) \geqslant 1 \tag{7b}$$

因此，在 EMH/RE 和贴现因子为常数的假设下，我们期望出现的情形是完全预见价格 P_t^* 的方差大于实际价格 P_t 的方差。为了表述方便，我们假定 P_t^* 由（3）式来进行计算。然而，许多实证研究都使用一种等价方法。DPV 公式（3）与欧拉方程是一致的。

$$P_t^* = \delta\ (P_{t+1}^* + D_{t+1}) \tag{8}$$

因此，如果假设终值为 P_{t+n}^*，那么我们就可以用（8）向后递归来算 P_t^*。事实上，这就是 Shiller（1981）使用的方法。对 Shiller（1981）的一个批评是他将股票价格的样本均值用作终值，即 $P_{t+n}^* = n^{-1} \sum_{t=1}^{n} P_t$。Marsh 和 Merton（1986）指出，这种用法会导致 P_t^* 估计的有偏性。然而，Grossman 和 Shiller（1981）及 Shiller（1989）则使用实际股票价格 P_{t+n} 来代替不可观测的 P_{t+n}^*。我们将可观测序列记为 P_{t+n}，在（8）中用 P_{t+n} 代替 P_{t+n}^*，这就保持了模型的逻辑结构，因为 $P_{t+n} = E\ (P_{t+n}^* \mid \Omega_{t+n})$。

考虑图 1 中的结果。图 1 使用的是 Shiller 关于美国 1871—1995 年期间的数据。点线表示（实际）股票价格 P_t，实际股票价格的波动程度显然大于完全预见股票价格 P_t^*，我们是在（实际）贴现率为常数的假设下来计算 P_t^* 的。“仔细考察”就会发现，图 1 意味着我们在常数（实际）收益假设下应拒绝 EMH - RVF（当然，我们不知道估计值的标准差为多大）。我们接下来看看进行更为严格的检验会得到什么结果。

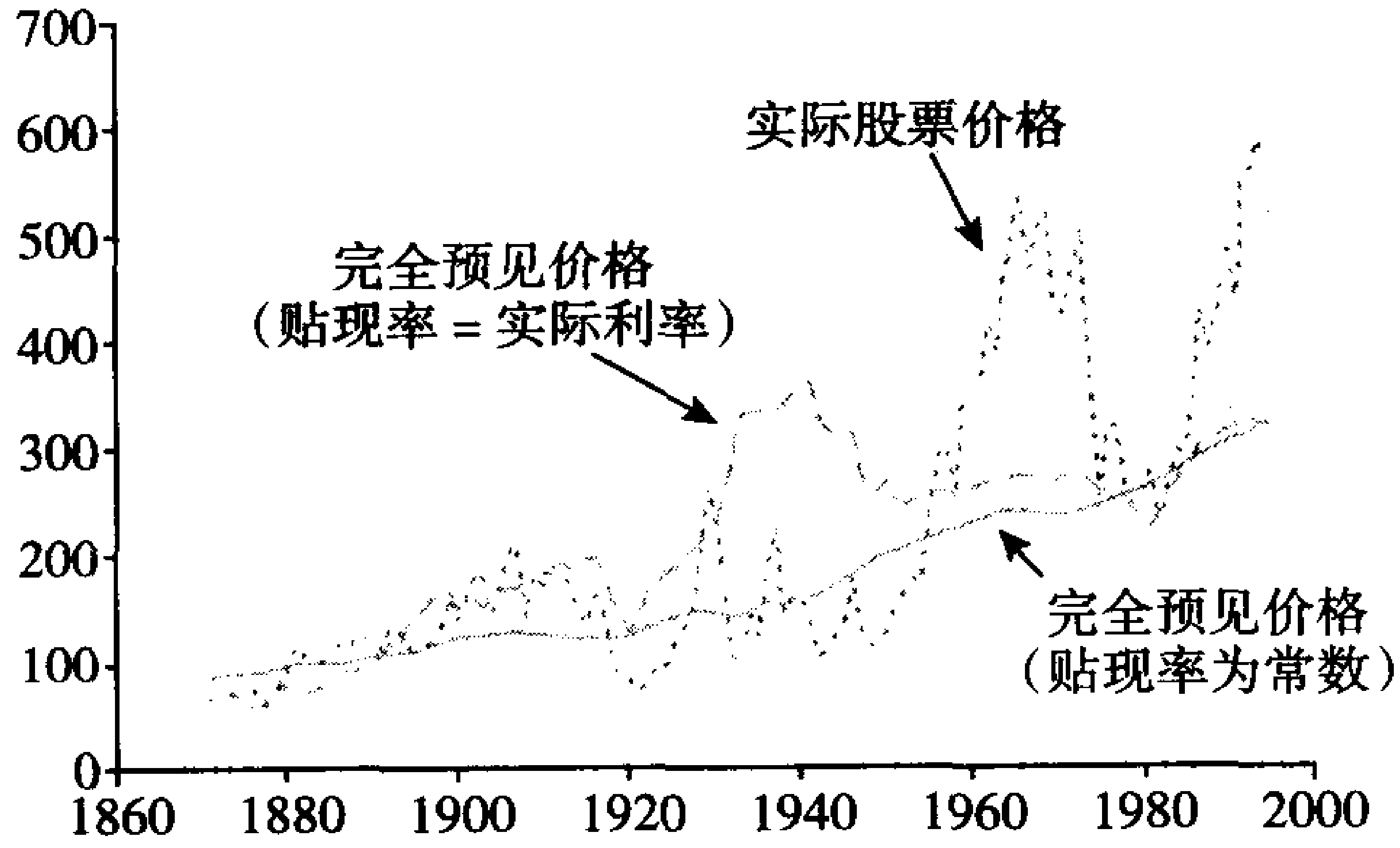


图 1　实际股票价格与完全预见股票价格（美国数据：1871—1995）

波动率检验

对理性定价公式（RVF）的实证检验普遍使用“实际变量”。对波动率检验的早期研究假设实际贴现率为常数，Shiller（1981）发现美国股票价格存在过度波动性，也就是说，不等式（7b）大致不成立（即 SDR = 5.59）。然而，LeRoy 和 Porter（1981）使用稍微不同的方法（见附录）却发现，尽管方差边界遭到了违反，但在边界上拒绝零假设却是统计显著的。

我们可以重新讨论完全预见价格，假设必要的实际收益为 k_t，因此，$\delta_t = (1+k_t)^{-1}$ 是时变的。例如，我们可以假设 k_t 等于实际利率加上常数风险溢价，即 $k_t = r_t + rp$。于是，kt_t 每年都会变化，P_t^* 的计算公式如下，

$$P_t^* = \frac{D_{t+1}}{(1+k_{t+1})} + \frac{D_{t+1}}{(1+k_{t+1})\ (1+k_{t+2})} + \cdots + \frac{(D_{t+n}+P_{t+n})}{(1+k_{t+1})\ \cdots\ (1+{}_{t+n})} \tag{9}$$

公式中的终值等于样本内实际价格的最后一个值。然而，尽管对时变实际利率进行了考虑，方差边界仍然遭到了违反（参见 Mankiw、Romer 和 Shapiro（1991），Scott（1990））。使用最新的美国数据就可以看出这一点（如图 1 所示）。就完全预见价格而言，在实际贴现率（等于实际利率）具有时变性时的价格变化比贴现率为常数时的价格变化大得多，但这里的情形是实际价格的波动远远高于完全预见价格的波动。如果我们将 1995—2004 年期间的数据包含进来，那么相对于两个完全预见价格序列而言，实际价格的波动率甚至在 20 世纪 90 年代就大得多。

我们可以从头开始进行上面的计算。在知道实际价格的变化后，我们可以计算实际收益的变化，实际收益必须使 Var（P_t^*）与 Var（P_t）相等。Shiller（1981）在 P_t 和 D_t 都有确定性趋势的假设下来进行计算，结果发现，实际收益的标准差每年必须大于 4%。然而，实际利率的历史变动每年小于 1%，这大大低于使方差边界检验有效所需要的变化。因此，对违背过度波动性的解释似乎并不意味着实际利率具有时变性。

为了“解决”方差边界遭到违反的问题，另一种思路是用以消费为基础的资本资产定价模型（CCAPM）来对时变风险溢价进行度量（见第 13 章）。对幂效用函数而言，时变贴现率取决于消费增长率 g、（常数）时间偏好率 θ 和相对风险规避系数 γ。为简化起见，假设红利增长率为常数，从而 $D_t = D_0 g^t$。因此，完全预见股票价格为

$$P_t^* = D_0 g^t \left[C_t^{\gamma} \sum_{s=0}^{\infty} (\theta g)^s C_{t+s}^{-\gamma} \right] \tag{10}$$

P_t^* 随着时间会发生变化（Grossman 和 Shiller（1981）），它取决于当前消费水平与未来消费 $C_{t+s}^{-\gamma}$ 加权平均值之间的相对关系。显然，这就使 P_t^* 比贴现率为常数假设下（即相对风险规避系数）的序列有着更大的波动性。

Shiller（1989）仍然使用 CCAPM 模型，但将原先按照常数进行增长的红利用实际红利来进行替代，取 $\gamma = 4$ 来对美国 1889—1985 年期间的方差边界检验重新进行了计算。P_t 和 P_t^* 的图形表明，直到 1950 年，方差边界检验都没有遭到违反。然而，从 1950 年以后，实际股票价格的波动与完全预见价格波动之间就不再呈现出近似关系。总的来说，以 CCAPM 模型（相对风险规避系数为常数，没有市场摩擦）为基础的时变贴现率假设似乎不能完全解释所观察到的股票价格运动。

统计问题

在这一节中，我们说明对 EMH - RVF 假设进行检验的困难和研究人员解决这些问题的技巧。从小的方面来看，这些研究为时间序列方面的应用计量提供了一些“案例研究”。

Flavin（1983）和 Kleidon（1986）指出 Var（P_t^*）和 Var（P_t）的度量存在小样本偏差，这使得“第一代”方差边界检验变得无效。Flavin（1983）在贴现率为常数的假设下对这些检验进行了分析并给出了两个要点。

首先，Var（P_t^*）和 Var（P_t）的小样本估计都存在向下的偏差，偏差大小取决于 P_t 和

P_t^* 之间序列相关的程度。因为 P_t^* 的自相关程度比强，所以 Var（P_t^*）估计值向下的偏差比 Var（P_t）大得多。因此，在有限样本中，即使 RVF 的零假设为真，Var（P_t^*）< Var（P_t）也可能出现。其次，Shiller 将样本平均值用作 $t+n$ 时的终点价格 P^*，这也会导致一定的偏差，这种偏差使得零假设遭到拒绝的可能性更大一些。

Gilles 和 LeRoy（1991）注意到与终端价格有关的一个问题。完全预见价格为

$$P_t^* = \sum_{i=1}^{\infty} \delta^i D_{t+i} \tag{11}$$

它是不可观测的。可观测序列 $P_{t|n}^*$ 可以用样本末的实际价格 P_{t+n} 来构造。然而，仍然存在的问题是，$P_{t|n}^*$ 的样本方差低估了 P_t^* 的真实（但无法观测的）方差。从直觉来看这是因为，$P_{t|n}^*$ “依赖于” P_{t+n}，$P_{t|n}^*$ 没有考虑样本外红利的变化（$P_{t|n}^*$ 一般隐含地将其设为零，而 P_t^* 则要考虑这些变化，因为和式是无穷级数求和）。显然，如果样本容量非常大（无穷大），那么这个问题的影响就会非常小，但在有限样本情形中，这个问题还是很重要。

Flavin（1983）也对“第一代检验”提出了批评，与这些检验的作者一样，Flavin 假设所使用的序列是平稳序列。后来的研究在价格和红利是非平稳序列（即存在一个随机趋势）的基础上来考察方差边界检验的有效性问题。直观地说，序列非平稳所导致的问题是，总体的方差是时间的函数，从而样本方差不是总体相应值的正确测度。然而，为了使方差边界检验变得有意义，如何“剔除”这些随机趋势还不是一件显而易见的事情。我们现在就转到这个问题的讨论上去。

11.2　波动率检验与平稳性

Shiller 的波动率不等式纯粹是基于如下假设的结果：实际股票价格是完全预见价格 P_t^* 的无偏最优估计

$$P_t^* = P_t + u_t \tag{12}$$

u_t 是随机误差项，$E(u_t \mid \Omega_t) = 0$。换句话说，P_t 是对 P_t^* 进行预测的充分统计量，除 P_t 之外的其他信息都无法改进对 P_t^* 的预测，P_t 在这种意义下是“最优的”，这意味着条件预测误差 $E[(P_t^* - P_t) \mid \Omega_t]$ 与 t 或之前可获得的信息独立。因为 $P_t \subset \Omega_t$，所以 P_t 独立于 u_t，Cov（P_t，u_t）= 0。根据平稳序列协方差的定义，由（12）就可以直接得到

$$\begin{aligned} \operatorname{Cov}(P_t^*, P_t) &= \operatorname{Cov}(P_t + u_t, P_t) \\ &= \operatorname{Cov}(P_t, P_t) + \operatorname{Cov}(P_t, u_t) \\ &= \sigma^2(P_t) \end{aligned} \tag{13}$$

因为 $\rho(P_t, P_t^*) = \operatorname{Cov}(P_t, P_t^*) / \sigma(P_t)\,\sigma(P_t^*)$，所以

$$\sigma(P_t) = \rho(P_t, P_t^*)\,\delta(P_t^*) \tag{14}$$

因为 ρ 的最大值为 1，所以（14）意味着如下熟悉的不等式成立：

$$\operatorname{Var}(P_t) \leqslant \operatorname{Var}(P_t^*) \tag{15}$$

平稳序列的总体均值、方差（标准差）和自协方差具有时间不变性（为常数）。对样本数据而言，运用（15）的困难在于我们不知道样本是否是从某个平稳序列总体中进行抽取的。

值得注意的是，（15）中的标准差是一个简单的无条件指标。如果时间序列图反映的结果是曲线经常改变方向、频繁地与均值相交（即成锯齿状），那么我们在短期样本数据中根据样本值就可以得到 $\sigma(P_t)$ 总体值的“良好”估计结果。然而，如果变量长时间缓慢地围绕均值运动，那么我们就需要一个很长时间的样本数据来得到“真实”总体方差的良好估计（我们所需要的代表性序列在样本数据中存在一定的周期性，而不是只包含 1/4 个周期或半个周期的数据）。事实上，股票价格运动具有持续性，从而股票价格长期运动存在一定的波动性（图

1），因此，即使数据序列是平稳的，为了对真实的标准差（方差）进行准确度量，我们需要一个长的样本数据集合。

如果序列非平稳，那么总体均值或方差具有时变性，（15）没有定义。因此，我们需要设计一个不同的方差不等式，方法是将 P_t 和 P_t^* 转换成“新的”平稳变量，这就产生了方差不等式的各种形式。问题在于，我们很难根据有限数据集合上的统计检验结果确定某个特定序列是否是平稳序列。例如，$x_t = x_{t-1} + \varepsilon_t$ 所生成的序列是非平稳序列，但 $x_t = 0.98x_{t-1} + \varepsilon_t$ 所生成的序列是平稳序列。然而，在（股票价格的）任意有限数据集合中通常很难在统计上对两者进行区分，因为在回归方程 $x_t = a + bx_{t-1} + \varepsilon_t$ 中“b”的估计值存在样本误差，通常将其取为 0.98 或 1。另外，$b = 1$ 的检验统计量分布不是“标准”分布，它的有限样本分布通常是未知的。

我们现在回到（实际）红利及由此产生的（对常数 δ 而言）P_t^* 和 P_t 是否为非平稳序列的问题上来。当我们根据有限样本数据对 σ（P_t）和 σ（P_t^*）进行估计时，由非平稳性所导致的差异到底有多大呢？例如，在一些早期工作中，Shiller（1981）使用简单的 $\lambda_t = e^{bt}$ 就将 P_t^* 和 P_t 中的趋势去除了，在整个样本期上对 $\ln P_t = a + bt$ 进行回归得到 b 的估计值。如果 P_t 具有随机趋势，那么假设定性趋势在统计上无效就可以将“趋势去掉”。由此产生的问题是，在 Shiller（1981）的研究中方差边界的违反是否应该归因于数据中的不合适趋势。

Kleidon（1986）及 LeRoy 和 Parke（1992）使用蒙特卡罗模拟（MCS）对方差边界进行了考察。例如，Kleidon（1986）假设预期红利增长率为常数，红利服从一个带漂移项的非平稳（几何）随机游走，飘移项为

$$\ln D_t = \theta + \ln D_{t-1} + \varepsilon_t \tag{16}$$

这里的 ε_t 是白噪声。根据（16）可以生成 D_t 序列，序列有 m 个观测值，使用 DPV 公式就可以生成 P_t^* 和 P_t 长度为 m 的时间序列，从而可以对 Var（$P_t > \text{Var}P_t^*$）进行考察。这是 MCS 的第一“轮”计算，重复“实验”m 次，看 Var（P_t）> Var（P_t^*）会出现多少次。因为根据构造，RVF 是“正确的”，所以在 m 次实验中，方差边界被违反的次数不可能很大。（偶然因素也可能导致方差边界被违反。）。事实上，Kleidon（1986）发现，如果使用的是生成数据（用 λ_t 来消除趋势），那么方差边界被违反的频率很高，频率高达 90%。而“大体上被违反”（即 VR > 5）的频率差异很大，频率取决于模拟中所假定的利率（贴现率）。（例如，当 $r = 7.5\%$ 时，“大体上被违反”的频率仅为 5% 左右，而当 $r = 5\%$ 时，这一频率上升到 40% 左右。）

Shiller（1989）对 Kleidon 的程序进行了精炼，他注意到 Kleidon 的研究将红利增长率假设与红利价格比率不合理值所隐含的利率合在一起进行使用。Shiller 对此进行了纠正，由此导致的结果是，在所生成的数据中，红利价格比等于历史平均值。在 RVF 的零假设下，他发现方差比大体上被违反的概率远远小于 Kleidon 中的概率。

Shiller（1989）还进一步发现，在那些使用模特卡罗模拟的上述研究中，违反方差不等式的频率都不没有 Shiller（1981）高，后者使用的是“真实世界”数据。Shiller 还指出，Kleidon（及其他人）使用的这个“特殊例子”（（实际）红利（的对数值）服从带漂移的随机游走模型）也许不能正确代表实际数据序列。

这一争论所强调的问题是结果是否可信，因为所使用的数据是在“特殊情形”中（如随机游走）用蒙特卡罗方法所生成的。这些蒙特卡罗研究对实证结果进行了非常特殊的“敏感性检验”，但“统计量”以特殊的参数化模型（如随机游走）为条件。人们可能会同意 Shiller（1989）的观点，在先前的经济环境中我们很难接受投资者的如下信念：如果面临的情形是无法预期的红利出现 $z\%$ 的增长，那么投资者会期望未来各期的红利出现 $z\%$ 的增长。然而，后者意味着（对数）红利服从随机游走模型。

上述争论的结果是，不仅方差边界检验的小样本性质不可靠，如果出现的是非平稳序列情形下，那么大样本下的检验也可能是错误的。显然，我们在实践中所能做的就是（随着“时间”的推移，等待新的时间序列出现）对不同去势方法所得到的波动率进行稳健性检验。例

如，Shiller（1989）对他早期的方差不等式重新进行了检验，这次使用的是 P_t/E_t^{30} 和 P_t^*/E_t^{30}，他用实际收益 30 年的移动平均值 E_t^{30} 来消除实际价格中的趋势。另外，他也使用了 P_t/D_{t-1} 和 P_t^*/D_{t-1}，D_{t-1} 是上一年的实际红利值。Shiller（1981）在整个样本上对定性趋势进行了估计，使用的信息是 t 时刻未知的信息。因此，Shiller（1989）的后续工作使用的是直到 t 时刻的数据来估计时间趋势，然后用这个时间趋势来消除 P_t^* 和 P_t 中的趋势（也就是说，$\lambda_t = \exp[b_t]t$，递归最小二乘法使用的数据越多，b_t 估计值就越有可能发生变化）。

使用 P_t^* 和 P_t 各种变形所得到的结果如表 1 所示。方差不等式（15）总是被违反，但违反程度没有 Shiller（1981）最初研究中的程度高，Shiller（1981）使用的是确定性趋势。然而，在所有变体中方差不等式（14）都被违反。

Shiller（1989）使用（对数）红利随机游走模型来重复 Kleidon 的蒙特卡罗模拟研究，用所生成的实际收益序列 E_t（假设所生成的收益与所生成的红利成比例）来去除 P_t^* 和 P_t 中的趋势，假设实际贴现率为 8.32% 的常数（等于样本中股票的年度平均实际收益）。在 1000 次蒙特卡罗模拟中，他发现 $\sigma(P_t/E_t^{30})$ 大于出现的频率为 75.8%。因为所生成的数据违反了所构造的基本模型，所以人们实际期望的结果是，在所生成的数据中 $\sigma(P_t/E_t^{30}) > \sigma(P_t^*/E_t^{30})$ 出现的概率应为 95%。因此，当红利不是平稳序列时，如果用来消除序列中的趋势，那么方差边界就有伪（spurious）违反的倾向。然而，对所生成的数据而言，VR 的均值等于 1.443，尽管大于 1.00 的“真实”值，但还是远低于实际数据中 4.16 的观测值。（在 1000 次模拟中，“所生成的”方差比仅有一次大于 4.16。）

表 1　　　　方差边界检验[1]

去势方法	ρ	$\sigma(P_t^*)$	$\rho\sigma(P_t^*)$	$\sigma(P_t)$[2]	VR
1　具有时变性的定性趋势	n. a	n. a	n. a	n. a	2.120
2　使用 D_{t-1}					
（a）常数贴现因子	0.133	4.703	0.62	6.03	1.28[3]
（b）时变贴现因子	0.06	7.779	0.47	6.03	1.29[3]
3　使用 E_t^{30}					
（a）常数贴现因子	0.296	1.611	0.47	6.706	3.77[3]
（b）时变贴现因子	0.048	4.65	0.22	6.706	1.44[3]

注：

1. 如果 RVF 成立，那么 $\sigma(P_t) = \rho(P_t, P_t^*)\sigma(P_t^*)$。
2. 在倒数第二列中，$\sigma(P_t)$ 是实际价格的标准差。
3. VR 是蒙特卡罗模拟均值，RVF 在这里成立。

资料来源：数字来自于 Shiller（1989），pp. 87 ~ 91。

显然，E_t^{30} 和 P_t^* 都是长期移动平均值，都具有长期平滑趋势，我们可以只用 P_t^*/E_t^{30} 在小样本中变化的一个部分。因此，样本方差可能会向下偏斜。因为 P_t 不是平滑的，所以（P_t/E_t^{30}）的波动率可能比 P_t^*/E_t^{30} 大，即使在更长的样本期中也会出现这种情形。这与如下事实矛盾：如果我们拥有长期数据，那么基于这些数据的统计检验只能是定性检验。表 1 所给出的蒙特卡罗模拟证据和结果对股票价格而言是与 EMH－RVF 相反的证据。

Mankiw、Romer 和 Shapiro（1991）对他们 1985 年的早期论文（Mankiw、Romer 和 Shapiro（1985））进行了改进，通过考虑 P 和 P^* 相对于天真预测（naive forecast，即下一年的预测值等于今年的值）P^o 的变化来处理非平稳问题。对天真预测而言，他们假设红利服从随机游走过程，因此 $E_tD_{t+j} = D_t$（对所有 j 都成立）。使用 RVF 就得到天真预测值

$$P_t^o = [\delta/(1-\delta)]D_t \tag{17}$$

其中，$\delta=1/(1+k)$，k 是股票的常数必要收益。现在考虑恒等式

$$P_t^* - P_t^o = (P_t^* - P_t) + (P_t - P_t^o) \tag{18}$$

RE 预测误差为 $P_t^* - P_t$，它与 $P_t - P_t^o$ 和 t 时刻的信息独立。(18) 除以 P_t，平方后再求数学期望就得到

$$q_t = E[(P_t^* - P_t^o)/P_t]^2 - \{E[(P_t^* - P_t)/P_t]^2 + E[(P_t - P_t^o)/P_t]^2\} \tag{19}$$

于是，不等式就变成

$$E[(P_t^* - P_t^o)/P_t]^2 \geqslant E[(P_t^* - P_t)/P_t]^2 \tag{20a}$$

$$E[(P_t^* - P_t^o)/P_t]^2 \geqslant E[(P_t - P_t^o)/P_t]^2 \tag{20b}$$

上面两个关系式的美妙之处在于，表达式中的每一项都是平稳的，用 P_t 进行平减后有可能使异方差问题的影响达到最小。(20a) 表明，与天真预测值相比市场价格是事后“理性价格” P_t^* 的更好预测．市场价格有着更小的均方误差。(20b) 表明，与市场价格相比事后理性价格 P_t^* 围绕天真预测值 P_t^o 的波动性要更大一些，它与 Shiller 波动率不等式类似。EMH - RVF 的另外一种检验方法假设如下方程中的 ψ 等于零：

$$q_t = \psi + \varepsilon_t \tag{21}$$

使用 1871—1988 年期间的年度数据和股票价格指数，Mankiw、Romer 和 Shapiro (1991) 发现，当必要收益为 $k=6$ 或 7% 时，我们在 5% 的显著性水平下拒绝 (21)（尽管在 5% 的必要收益假设下模型遭到拒绝的可能性更大一些）。当 Mankiw、Romer 和 Shapiro 将必要均衡名义收益设为（名义）无风险利率加常数风险溢价（即 $k_t = r_t + rp$）时，与常数实际收益的情形相比，使用 (19) 时的 EMH - RVF 遭到拒绝的可能性要更大一些。

Mankiw、Romer 和 Shapiro (1991) 还考虑了 Fama 和 French (1988b) 所使用的回归类型的检验。更为特别的是，他们考虑了收益的如下自回归方程：

$$[(P_t^{*n} - P_t)/P_t] = \alpha + \beta[(P_t - P_t^o)/P_t] + u_t \tag{22}$$

这里的 P_t^{*n} 是某个特定时期 ($n=1, 2, \cdots$) 上的完全预见价格。Mankiw、Romer 和 Shapiro (1991) 使用蒙特卡罗研究后发现：如果一些合理的条件在真实数据中成立，那么 β 的估计值和标准误差存在非常严重的小样本偏差问题，且偏差随着 n 的增大而增大。然而，当他们使用年度数据时，在常数实际收益率（每年 5%、6% 或 7%）条件下，就 1 到 10 年中的绝大多数期限而言仍然出现的情形是在 1% ~5% 的显著性水平拒绝零假设 $H_0: \beta=0$。如果是在均衡收益等于名义利率加常数风险溢价的假设下来构建 P_t^*，那么对五年以上的期限而言，零假设 $H_0: \beta=0$ 仅在 5% ~10% 的显著性水平下遭到拒绝。Mankiw 等的研究结果表明，当我们对检验统计量的小样本性质进行仔细考察时，Fama - French 所发现的“长期收益（即 $n>5$ 年）具有可预测性”的证据不是很明显。

Mankiw、Romer 和 Shapiro (1991) 的论文还对方差边界检验解释中所遇到的另外一个问题（在计算完全预见价格 P_{t+n} 时最终价格的重要性问题）进行了解决。Merton (1987) 指出，样本终点价格 P_{t+n} 使得样本外事件对（样本期内的）股票价格有影响，因为它反映的是（不可观测的）未来红利。因此，如果 P_{t+n} 不能很好地代表 $E_t P_{t+n}^*$，那么使用实际价格的最后一个值来进行波动率检验时存在度量误差。但 Mankiw 等和 Shiller (1989) 指出，如果“样本内”的实际红利支付足够的高以至于样本外事件（用 P_{t+n} 来进行度量）的重要性受到了限制，那么 Merton 的批评就不是很重要。从实证角度来看，Mankiw、Romer 和 Shapiro (1991) 所使用的数据就是后面一种情形，数据的样本期很长。然而，他们对这一争论提出了一种巧妙而又简单的权衡方法（也可参见 Shea (1989)）。对不同时期而言 ($n=1, 2, ...$)，使用方程 (3) 就可以计算出完全预见股票价格，也可以计算 q_t 出的值。因此，使用给定样本数据集合的最后一个价格就可以计算出许多 q_t^n ($n=1, 2, ...$) 的值。因此，他们根本不用担心样本期的最后一个价格会对结果起主导作用。一般而言，他们发现有效市场假说成立的证据短期 ($n=1$ ~5 年) 比长期 ($n>10$ 年) 更有力。

在后续研究中，Gilles 和 LeRoy (1992) 假设红利价格比序列是平稳的，方差不等式为

$\sigma^2(P_t \mid D_t) \leqslant \sigma^2(P_t^* \mid D_t)$。方差的样本值（与 Shiller（1981）一样，使用美国 1871—1988 年的股票价格指数数据）意味着存在过度波动率，因为 $\sigma^2(P_t \mid D_t)=26.4$，$\sigma^2(P_t^* \mid D_t)=19.4$。然而，他们也注意到 $\sigma^2(P_t^* \mid D_t)$ 的样本方差存在向下偏误的两个原因：首先，$(P_t^* \mid D_t)$ 是序列正相关的（Flavin（1983））；其次，在终端时刻，我们的假设是不可观测的 $E_tP_{t+n}^*$ 等于实际（终端）价格 P_{t+n}。他们使用蒙特卡罗模拟实验发现偏差的第一个来源最为重要，也最为严重。蒙特卡罗实验假设红利服从随机游走过程，$\sigma^2(P_t^* \mid D_t)$ 在蒙特卡罗模拟运算中的值为 89.3，而在实际样本数据中的值为 19.4。另一方面，$\sigma^2(P_t \mid D_t)$ 的样本值是总体方差相当精确的一种度量。因此，Gilles 和 LeRoy 的结论是，Shiller 类型的方差边界检验“无法定论”（indecisive），但也不是所有检验都没有用。Gilles 和 LeRoy 在 P_t 和 P_t^* 正交性的基础上发展了一种检验方法（West（1988）），这种方法的稳健性更好一些。这种“正交性检验”使用（对数）红利的随机游走假设，与 Shiller 类型检验统计量相比，“正交性检验”所涉及的检验统计量样本偏差和波动程度都要小一些。正交性检验非常明确地拒绝现值模型（要注意的是，这一程序涉及一些细微差别，我们不再赘述）。因此，Gilles 和 LeRoy 研究的合理结论是，在（对数）红利随机游走模型假设下要拒绝 RVF。

Scott（1990）使用一种稍微不同的程序来对如下方程进行估计：

$$P_t^* = a + bP_t + \varepsilon_t \tag{23}$$

其中，EMH - RVF 意味着 $a=0$ 且 $b=1$。Scott 用上一年的红利值来对 P_t 和 P_t^* 进行平减，目的是为了使变量变成平稳变量，他们还对回归残差中的序列相关性进行了“纠正”。结果发现，对美国股票价格数据而言，上述限制条件被违反了，从而 P_t 不是 P_t^* 的无偏预测。回归的 $\bar{R}^2$ 非常低，因此 P_t 和 P_t^* 之间几乎不存在（正）相关关系，P_t 是事后完全预见价格非常糟糕的预测（然而需要注意的是，EMH - RVF 的拒绝而言，重要的是统计量的无偏性而不是 $\bar{R}^2$）。然而，有效市场假说确实意味着（23）中所包含的任何信息 Ω_t 在统计上都不显著。Scott（1990）将 $(P_t^* - P_t)$ 对红利价格比（即红利率）进行回归，结果表明红利价格比变量在统计上显著，因此我们应拒绝信息效率假设。

Shiller（1989）进行了类似的回归，他用 E_t^{30} 来对价格进行平减：

$$(P_t^*/E_t^{30}) = a + b\,(P_t/E_t^{30}) + \varepsilon_t$$

结果发现 $\hat{b} \ll 1$。使用蒙特卡罗模拟所发现的结果是 $\hat{b}$ 存在向下的偏差，但在使用实际数据时这个偏差不足以显著地拒绝 RVF。

11.3　比索问题与方差边界检验

当存在“比索问题”（Peso problem）时，即使有效市场 - 理性定价公式（EMH - RVF）成立，在样本数据中也可能出现方差边界遭到违反的情形。假设在我们所拥有的样本数据中投资者认为将来会获得高红利的概率很小，投资者的红利预期是“高”红利与“正常”红利的加权平均。然而，我们假设“高”红利从来没有发生过。于是，样本外的实际红利 D_{t+j} 低于投资者的真实预期（即 $D_{t+j} < E_{t+j-1}D_{t+j}$）。投资者在样本期内存在系统预测误差。将样本期进行扩展也会发现投资者预期更低红利（从未发生过）的情形，因此 $D_{t+j} > E_{t+j-1}D_{t+j}$。在扩展的“整个”样本期上，平均预测误差为零。我们只能对第一个样本期中的数据进行“观察”，而所观察到的红利值是预期红利的一个有偏度量，所以比索问题产生。为了说明比索问题，考虑一种资产，它的红利支付现金流为 D_{t+j}，所有现金流按照固定贴现率 δ 进行贴现。为了简化符号，我们将时期“$t+1$”想象为包含 m 个不同的时期（$t=0, 1, 2, \cdots, m$）。在理性定价公式中，股票价格为

$$P_t = \delta E_t D_{t+1} \tag{24}$$

这里的 δ 为常数贴现因子。假设 $1-\pi$ 为状态 1 发生（即“正常红利”）的概率，因此投资者的真实预期为：

$$\begin{aligned} E_tD_{t+1} &= \pi E_t\ (D_{t+1} \mid Z2) + (1-\pi)\ E_t\ (D_{t+1} \mid Z1) \\ &= \pi\ [E_t\ (D_{t+1} \mid Z2) - E_t\ (D_{t+1} \mid Z1)] + E_t\ (D_{t+1} \mid Z1) \\ &= \pi \nabla D_{t+1} + E_t\ (D_{t+1} \mid Z1) \end{aligned} \tag{25}$$

这里的 $E_t\ (D_{t+1} \mid Z1)$ 表示在状态 1 下对“正常”红利的预期，$E_t\ (D_{t+1} \mid Z2)$ 表示在状态 2 下对“高”红利的预期，我们假设 $E_t\ (D_{t+1} \mid Z2) > E_t\ (D_{t+1} \mid Z1)$。为了进一步简化问题，假设状态 1 下红利预期值等于常数且事后值等于 D，于是 $E_t\ (D_{t+1} \mid Z1) = D$。比索问题的关键在于，研究人员只拥有“高利润”并没有实现的那些时期中的数据（在 $t=0,\ 1,\ 2,\ \cdots,\ m$），但高利润存在的可能性确实对投资者的“真实”预期有影响，对股票价格也有影响。理性投资者认为股票价格应为

$$\begin{aligned} P_t &= \delta E_tD_{t+1} \\ &= \delta\ [\pi \nabla D_{t+1} + E_t\ (D_{t+1} \mid Z1)] \\ &= \delta\pi\ [E_t\ (D_{t+1} \mid Z2) - D] + \delta D \end{aligned} \tag{26}$$

我们在这里假设高利润没有发生，红利实际上仍然为常数值 D。由（26）给出的实际价格 P_t 可能发生变化，原因要么是投资者对 π 的看法发生了改变，要么是对高利润情形下未来红利 $E_t\ (D_{t+1} \mid Z2)$ 的看法发生了改变，于是 $\mathrm{Var}\ (P_t) > 0$。但是，如果高利润永远不发生，那么投资者将获得的红利为常数 D，研究人员将事后完全预见价格用常数值 $P_t^* = \delta D$（$t=0,\ 1,\ 2,\ \cdots,\ m$）来进行度量，于是 $\mathrm{Var}\ (P_t^*) = 0$。因此，这就违反了方差边界假设，即 $\mathrm{Var}\ (P_t) > \mathrm{Var}\ (P_t^*)$，尽管价格始终等于由（26）所给出的基础价值。这是一个样本数据结果，并不代表总体的行为。如果数据集合的时间越长，那么预期事件就可能发生，从而 P_t^* 随着实际价格而发生变化。

考虑比索问题的含义。回归方程是将 P_t^* 对 P_t 进行回归：

$$P_t^* = \alpha + \beta P_t + \omega_t \tag{27}$$

如果不存在比索问题，那么我们预期的结果是 $\alpha = 0$ 且 $\beta = 1$。假设“正常红利” $E_t\ (D_{t+1} \mid Z1)$ 随着时间而发生变化，而高利润却没有发生过，研究人员用 $P_t^* = \delta D_{t+1}^{(1)}$（$t=0,\ 1,\ 2,\ \cdots,\ m$）来度量完全预见价格，这里的 $D_{t+1}^{(1)}$ 是红利在状态 1 下的下一期值，而实际价格为

$$P_t = \delta E_tD_{t+1} = \delta\ [\pi \nabla D_{t+1} + E_t\ (D_{t+1} \mid Z1)] \tag{28}$$

将 P_t^* 对 P_t 进行回归

$$(\delta D_{t+1}^{(1)}) = a + b\ \{\delta\ [\pi \nabla D_{t+1} + E_t\ (D_{t+1} \mid Z1)]\} + \omega_t \tag{29}$$

只有当预期不发生状态转移时（即 $\nabla D_{t+1} = 0$）才会有 $b=1$，此时 $E_t\ (D_{t+1} \mid Z1) = D_{t+1}^{(1)} - \varepsilon_{t+1}$ 是下期红利的无偏预测。如果 π 是时变的，那么这会导致 b 的估计是有偏的，因为这会导致实际价格 P_t 的系统变化与完全预见价格 $P_t^* = (\delta D_{t+1}^{(1)})$ 变化不匹配。

比索问题的产生是因为，一次性（one-off）的“特定事件”在样本期内该发生但实际上却没有发生。比索问题使得 EMH-RE 的假设检验变得非常复杂，EMH-RE 假设样本外数据与给予（零均值）随机误差的预期值不同。

11.4 波动率检验与回归检验

考察波动率检验与回归检验之间关系最简单的方法是观察如下回归方程：

$$P_t^* = a + bP_t + \varepsilon_t \tag{30}$$

系数 b 为

$$b = \mathrm{Cov}\ (P_t,\ P_t^*)\ /\sigma^2\ (P_t)$$

$$=\rho\ (P_t,\ P_t^*)\ \sigma\ (P_t^*)\ /\sigma\ (P_t) \tag{31}$$

将方差等式（14）中的ρ代入（31）后我们会看到，如果方差不等式成立，那么回归方程（30）中的等于1。现在考虑

$$P_t^* = a + bP_t + c\Omega_t + \eta_t \tag{32}$$

在EMH－RVF下我们预期H_0: $a=c=0$，$b=1$。如果确实如此，那么（31）可以简化为

$$P_t^* = P_t + \eta_t \tag{33}$$

因此方差边界检验也必须成立。零假设下的两种检验等价。稍微做一些变形，考虑$c=0$而$b<1$的情形（前面的很多实证研究都是这种情形），于是（32）可以简化为

$$\mathrm{Var}\ (P_t^*)\ = b^2\mathrm{Var}\ (P_t)\ + \mathrm{Var}\ (\eta_t) \tag{34a}$$

$$\mathrm{Var}\ (P_t^*)\ - \mathrm{Var}\ (P_t)\ =\ (b^2-1)\ \mathrm{Var}\ (P_t)\ + \mathrm{Var}\ (\eta_t) \tag{34b}$$

理性预期假设意味着Cov（P_t，η_t）$=0$。如果$b^2<1$，那么（34b）右边第一项为负；也可能出现的情形是，（34b）右边整个部分都为负。因此，$b<1$也意味着方差边界检验不成立。接下来考虑Fama和French（1988b）的长期回归方程：

$$h_{t,t+k} = a + bh_{t-k,t} + \eta_{t+k} \tag{35}$$

$$p_{t+k} = a +\ (b+1)\ p_t - bp_{t-k} + \eta_{t+k} \tag{36}$$

我们在这里使用了恒等式$h_{t,t+k} \equiv p_{t+k} - p_t$，在常数预期收益的零假设下，我们有$a\neq 0$且$b=0$，因此

$$p_{t+k} = a + p_t + \eta_{t+k} \tag{37}$$

对$b=0$而言，Fama－French回归与股票价格的随机游走模型一致。

我们已经证明了在市场有效的零假设下使用P_t和P_t^*的回归检验与方差边界不等式和基于长期股票收益的自回归模型是一致的。

11.5 小结

* Shiller（1981）最先使用方差边界不等式来进行检验，结论是果断地拒绝理性定价公式。在20世纪80年代的研究工作中，一些研究人员指出了Shiller方法的不足之处（如Kleidon（1986）、Flavin（1983）），但Shiller（1989）对这些批评作了非常成功的回应。

* 后续工作（如Mankiw、Romer和Shapiro（1991））及Gilles和LeRoy（1992）很肯定地指出理性定价公式的违反在统计意义上并不显著，对此问题的判定需要一种权衡观点。这些作者所引用的证据还不足以拒绝EMH。

* Shiller波动率不等式的直观吸引力和方法背后的简洁性和优美性被实际（统计）检验中的一些问题所遮蔽，尤其是平稳性问题和小样本偏差问题。计量经济学的最新发展为非平稳性问题的处理和时变风险溢价的建模提供了更为满意的方法，时变风险溢价的建模在下一章中进行讨论。

附录：LeRoy－Porter检验和West检验

前面所介绍的检验方法不能很好地适应本章的主体内容，但对这一领域的文献而言，这些检验都具有里程碑意义。因此，我们在这里对这些检验及相互关系进行讨论。

LeRoy和Porter（1981）的方差边界检验是基于这样一个数学性质：任何随机变量的条件期望比变量本身的波动性更小。他们的分析始于对未来红利的预测，预测的基础有限信息集$\Lambda_t=$（D_t，D_{t-1}）。基于Λ_t的未来红利预测值为

$$\hat{P}_t = E\ (P_t^* \mid \Lambda_t)\ = \sum_{i=1}^{\infty} \delta^i E\ (D_{t+i} \mid \Lambda_t) \tag{A1}$$

实际股票价格 P_t 的预测由整个信息集 Ω_t 来确定

$$P_t = E\ (P_t^* \mid \Omega_t) \tag{A2}$$

对（A2）运用迭代期望法则就得到

$$E\ (P_t \mid \Lambda_t)\ = E\ [E\ (P_t^* \mid \Omega_t)\ \mid \Lambda_t]\ = E\ (P_t^* \mid \Lambda_t) \tag{A3}$$

应用（A1）和（A3）后就得到

$$\hat{P}_t = E\ (P_t \mid \Lambda_t) \tag{A4}$$

因为 $\hat{P}_t$ 是 P_t 的条件期望，所以根据（A4），LeRoy - Porter 方差不等式为

$$\mathrm{Var}\ (\hat{P}_t)\ \leqslant \mathrm{Var}\ (P_t) \tag{A5}$$

假设变量是平稳的，（A5）中给出的 $\hat{P}_t$ 和 P_t 样本方差值是总体值的一致估计。如果 δ 已知且刻画 D_t 的 ARMA 模型已给定，那么 P_t 序列就可以由（A1）来构造。与 Shiller（1981）一样，LeRoy - Porter 的方法也能得到一个方差不等式。然而，LeRoy - Porter 检验也导致了某种形式的“正交性检验”。为了看清这一点，我们定义单期预测误差：

$$e_{t+1} \equiv\ (D_{t+1} + P_{t+1})\ - E\ (D_{t+1} + P_{t+1} \mid \Omega_t) \tag{A6}$$

欧拉方程为

$$P_t = \delta E\ (P_{t+1} + D_{t+1} \mid \Omega_t) \tag{A7}$$

将（A6）代入（A7）并向前迭代就得到

$$P_t^* = P_t + \sum_{i=1}^{\infty} \delta^i e_{t+i} \tag{A8}$$

其中，$P_t^* = \sum_{i=1}^{\infty} \delta^i D_{t+i}$。如果 e_t 是平稳的，那么

$$\mathrm{Var}\ (P_t^*)\ = \mathrm{Var}\ (P_t)\ + 2\mathrm{Cov}\ (P_t,\ \sum \delta^i e_{t+i})\ +\ [\delta^2/\ (1-\delta^2)]\ \mathrm{Var}\ (e_t) \tag{A9}$$

在理性预期假设下，e_{t+i}彼此不相关且协方差项为零，于是

$$\mathrm{Var}\ (P_t^*)\ = \mathrm{Var}\ (P_t)\ +\ [\delta^2/\ (1-\delta^2)]\ \mathrm{Var}\ (e_t) \tag{A10}$$

方程（A10）既是一个方差等式，又是 RVF 的一个正交性检验。在这个结合处注意方程（A9）与 Shiller 不等式一致是值得的。如果我们通过数据发现 Shiller 不等式遭到了违反，即 $\mathrm{Var}(P_t) \geqslant \mathrm{Var}\ (P_t^*)$，那么根据（A9）可知，这意味着 $\mathrm{Cov}\ (P_t,\ \sum \delta^i e_{t+i})\ < 0$。因此，单期预测误差的加权平均值与 t 时刻的信息（即 P_t）有关。因此，对 Shiller 方差边界的违反意味着（加权平均）单期收益可预测。于是，我们就得到了 Shiller 方差边界检验与单期收益可预测性之间的联系。

West（1988）的检验非常重要，因为即使红利是非平稳序列（但与股票价格是协整的），检验仍然有效，这种检验不需要用某个变量来替换不可观测的 P_t^*。与 LeRoy - Porter 方差不等式一样，West 检验基于数学期望的一个特殊性质，即受限信息集 Λ_t 上预测误差的方差肯定大于完整信息集合 Ω_t 上预测误差的方差。可以证明，West 不等式隐含着 LeRoy - Porter 不等式。我们还是从（A1）开始，注意

$$\delta\ (\hat{P}_{t+1} + D_{t+1})\ = \delta D_{t+1} + \delta E\ (\delta D_{t+2} + \delta^2 D_{t+3} + \cdots \mid \Lambda_{t+1}) \tag{A11}$$

对（A11）使用迭代期望法则 $E\ [E\ (.\ \mid \Lambda_{t+1})\ \mid \Lambda_t]\ = E\ (.\ \mid \Lambda_t)$ 后就得到

$$\delta E\ (\hat{P}_{t+1} + D_{t+1} \mid \Lambda_t)\ = \sum_{i=1}^{\infty} \delta^i E\ (D_{t+i} \mid \Lambda_t) \tag{A12}$$

将（A12）的左边代入（A1），

$$E\ (\hat{P}_{t+1} + D_{t+1} \mid \Lambda_t)\ = \delta^{-1} \hat{P}_t \tag{A13}$$

现在定义收益基于信息集合 Λ_t 的预测误差：

$$\hat{e}_{t+1} \equiv \hat{P}_{t+1} + D_{t+1} - E\ (\hat{P}_{t+1} + D_{t+1} \mid \Lambda_t) \tag{A14}$$

将（A13）代入（A14），

$$\hat{e}_{t+1} = P_{t+1} + D_{t+1} - \delta^{-1}\hat{P}_t \qquad (A15)$$

方程（A15）是基于 Λ_t 的预测误差，（A6）是基于 Ω_t（$\Lambda_t \subset \Omega_t$）的预测误差，根据条件数学期望法则，有

$$\mathrm{Var}\ (\hat{e}_{t+1}) > \mathrm{Var}\ (e_{t+1}) \qquad (A16)$$

（A16）就是 West 不等式，West（1988）使用美国数据发现该不等式不成立。尽管红利序列是非平稳的，（A16）的总体方差还是一个常数（平稳的），样本方差是总体方差的一致估计。然而不幸的是，只有红利序列非平稳（如 $D_{t+1} = D_t + \omega_{t+1}$）时上述结论才成立。LeRoy 和 Parke（1992）在红利服从几何随机游走模型（$\ln D_{t+1} = \ln D_t + \varepsilon_{t+1}$）的假设对 West 检验的性质进行了考察，读者可以参阅他们的论文以便获得更多信息。

LeRoy - Porter 不等式和 West 不等式都可以通过考虑受限信息集合和完整信息集合下的方差来导出。因此，你可能会猜想他们在 RVF 有效的前提下所作出的推断是一致的。事实确实如此，正如我们现在就要说明的那样。LeRoy - Porter 不等式（A10）对任意信息集合都成立，因此对 Λ_t 也成立，这意味着

$$\mathrm{Var}\ (P_t^*) = \mathrm{Var}\ (\hat{P}_t) + \frac{\delta^2}{(1-\delta^2)}\mathrm{Var}\ (\hat{e}_t) \qquad (A17)$$

而 LeRoy - Porter 不等式（A5）为

$$\mathrm{Var}\ (\hat{P}_t) \leqslant \mathrm{Var}\ (P_t)$$

将（A17）的 Var（$\hat{P}_t$）和（A10）的 Var（P_t）代入上式就得到了 West 不等式（A16）。

第12章

股票价格：向量自回归方法

目的

- 使用向量自回归（Var）方法来发展股票理性定价公式（RVF）的检验方法并给出一些例子。
- 说明使用向量自回归（Var）的检验、Shiller 波动率检验和 Fama－French 长期收益检验之间的关系。
- 说明尽管单期收益很难预测，但长期价格会明显偏离基础价值，这就产生了股票价格的过度波动性。
- 考察股票价格波动率与单期收益持续性之间的关系。

我们在第 11 章已经看到，股票价格的方差边界检验存在多种类型，在对这些检验结果进行解释时，变量平稳性是一个重要的问题，也就是说，我们要用一些合适的方法来“去除”实际价格 P_t 和 P_t^* 完全预见价格序列中的趋势。向量自回归（Var）方法是一种直接处理变量平稳性问题的方法，这种方法可以对检验 RVF 有效性的几种度量指标同时进行考虑。

理性定价公式 RVF 是必要收益（或贴现因子）的非线性函数，对其进行线性化处理是可能的。将向量自回归（Var）方法应用于股票价格 p_t' 时，对“理论”（对数）价格和实际（对数）价格 p_t 进行比较就可以确定实际价格是否存在过度波动性。因此，我们可以将使用 Var 方法所得到的结果与 Shiller 的波动率检验结果进行比较。注意，在 Fama－French（1988b）和其他研究中，长期收益（如 3～5 年）比短期收益（如 1 个月或 1～3 年）的可预测性更强。使用理性定价公式的线性化形式就可以导出长期收益公式，向量自回归方法为Fama－French 研究提供了一些补充证据。

理性定价公式假设股票价格只是在“基本面”方面的新信息或消息到达时才会发生变化：也就是说，当未来红利或贴现率方面出现新信息或消息时股票价格才会发生变化。一个感兴趣的问题是，在所观测到的股票价格波动运动中，多大比例可归因于收益或红利方面的消息。回答这个问题的关键是单期收益是否具有持续性。这里所说的“持续性”是指当前收益方面的消息对未来各期收益和贴现率都有比较大的影响。如果持续性很强，我们可以证明即使单期收益很难预测，收益方面的消息对股票价格也有很大的影响。这一章在对时变风险溢价的来源进行考察时，我们将对这一主题进行深入发展。

12.1 收益与理性定价公式的线性化

我们以理性定价公式的概述作为讨论的开始，根据红利价格比对理性定价公式进行重新

表述，因为红利价格比是将 Var 方法应用于金融市场时的一个关键变量。然后，我们导出理性定价公式和理性预期假设所隐含的 Wald 约束，证明这些约束条件意味着单期超额收益不可预测。

我们将 P_t 定义为 t 期末的股票价格，D_{t+1} 是 $t+1$ 期的红利支付，H_{t+1} 是 t 期末到 $t+1$ 期末的单期持有期收益。所有变量都取实际值。将 h_{t+1} 定义为

$$h_{t+1} \equiv \ln\ (1+H_{t+1})\ = \ln\ [\ (P_{t+1}+D_{t+1})\ /P_t] \tag{1}$$

单期收益与资本利得（P_{t+1}/P_t）和红利收益（D_{t+1}/P_t）正相关，对方程（1）进行线性化后就得到（见附录）

$$h_{t+1} \approx \rho p_{t+1} - p_t + \ (1-\rho)\ d_{t+1} + k \tag{2}$$

小写字母表示变量的对数值（如 $p_t = \ln P_t$），$\rho = \bar{P}/\ (\bar{P}+\bar{D})$ 是线性化参数，对年度数据而言，实证研究计算的结果通常为 0.94 左右，而 k 是线性化常量（就我们的目的而言，可以将其忽略）。方程（2）是一种逼近，我们将其视为精确逼近。（对数）红利价格比为

$$\delta_t = d_t - p_t \tag{3}$$

方程（2）变成

$$h_{t+1} = \delta_t - \rho\delta_{t+1} + \Delta d_{t+1} + k \tag{4}$$

对单期收益而言，毫无疑问，方程（4）看上去有些奇怪，也不是很直观。到目前为止，它是一个没有经济意义的（近似）恒等式。这意味着，如果要预测单期收益，那么我们需要预测未来的红利价格比 δ_{t+1} 和 $t+1$ 期的红利增长率。细心的读者可能会注意到，方程（4）是 δ_t 的前向差分方程，因为 $\delta_t = d_t - p_t$，向前求解就可以得到股票（对数）价格的表达式，因此可以将这个公式视为理性定价公式的对数表示。

使用向前递归的通常方法来求解方程（4）（施加横截性条件）（见 Campbell 和 Shiller (1988)）

$$\delta_t = \sum_{j=0}^{\infty} \rho^j\ (h_{t+j+1} - \Delta d_{t+j+1})\ \ - k/\ (1-\rho) \tag{5}$$

方程（5）也是一个恒等式（线性近似）。我们现在为方程（4）和（5）注入经济含义。在有效市场假设下，如果投资者愿意持有股票，投资者所要求的（对数）预期（实际）收益记为 r^e_{t+j}，于是

$$E_t h_{t+j} = r^e_{t+j} \tag{6}$$

根据方程（4），

$$\delta_t - \rho\delta^e_{t+1} + \Delta d^e_{t+1} + k = r^e_{t+1} \tag{7}$$

（注意 δ_{t+1} 的上标 e。）向前求解方程（7），

$$\delta_t = \sum_{j=0}^{\infty} \rho^j\ (r^e_{t+j+1} - \Delta d^e_{t+j+1})\ \ - k/\ (1-\rho) \tag{8}$$

对恒等式（4）取期望并进行迭代求解就可以得到几乎差不多的结果

$$\delta_t = \sum_{j=0}^{\infty} \rho^j\ (h^e_{t+j+1} - \Delta d^e_{t+j+1})\ \ - k/\ (1-\rho) \tag{9}$$

因为 δ_t 在 t 时刻是已知的，所以 $E_t\delta_t = \delta_t$。这里的唯一差异是 h^e_{t+j} 必须理解为股票的单期必要期望收益。在接下来要讨论的实证研究中，许多分析都着眼于使用（2）式右边“基础变量”（单期收益和红利增长率）的 *Var* 预测方程，使用方程（9）来给出价格红利比的预测值，我们将其记为 δ'_t，所得的 δ'_t 序列可以用来和实际值 δ_t 进行比较。使用恒等式 $p'_t = d_t - \delta'_t$ 就可以得到（对数）股票价格的预测值，也可以将股票价格的预测值与实际股票价格 p_t 的变动进行比较。

方程（9）是 Gordon 增长模型 $D/P = r - g$ 的一个动态版本，模型中的必要收益和红利增长率在每一期中都有变化。因此，现在的问题就变成对如下问题进行考察：当我们考虑时变红利和时变贴现率时，RVF 是否成立。

股票价格与向量自回归方法

尽管在用实际（而不是名义）数据进行检验时，股票价格和红利都可能是非平稳序列，但红利价格比和变量（$r_{t+1}-\Delta d_{t+1}$）更有可能是平稳序列，因此标准统计结果适用于方程（9）。如果理性定价公式成立，那么我们预期 δ_t 是（$r_{t+j}-\Delta d_{t+j}$）的 Granger 原因，因为理性定价公式（9）是未来变量的线性函数，我们可以用 Var 来预测未来的均衡收益和红利增长率。所选取的代表性投资者的信息变量向量为：

$$z_t = (\delta_t,\ rd_t)' \tag{10}$$

其中，$rd_t \equiv r_t - \Delta d_t$。为了说明问题，Var 模型的滞后长度取为 1，我们有

$$z_{t+1} = \begin{pmatrix} \delta_{t+1} \\ rd_{t+1} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \delta_t \\ rd_t \end{pmatrix} + \begin{pmatrix} w_{1,t+1} \\ {}_{2,t+1} \end{pmatrix} \tag{11}$$

$$z_{t+1} = Az_t + w_{t+1}$$

如果误差项没有异方差，那么可以用 GMM 方法来对误差进行修正，从而使用最小二乘法来估计系数 a_{ij}。在 Var 中增加滞后变量是为了确保残差序列不相关。滞后阶数很大的 Var 通常可以写成一阶系统（伴随形式），所以对一阶滞后 Var 的分析没有任何约束。定义 $e1' = (1,\ 0)$，$e2' = (0,\ 1)$，于是

$$\delta_t = e1'z_t \tag{12a}$$

$$E_t rd_{t+j} = e2'E_t z_{t+j} = e2'\ (A^j z_t) \tag{12b}$$

将（12）代入（8）就得到

$$e1'z_t = \sum_{j=0}^{\infty} \rho^j e2'A^{j+1}z_t = e2'A\ (1-\rho A)^{-1}z_t \tag{13}$$

如果（13）对所有 z_t 都成立，那么非线性约束（忽略常数项，因为所有数据都是相对于均值的偏差来进行度量的忽略常数项约束一阶系统增长率作比较）为

$$f\ (a)\ = e1' - e2'A\ (I-\rho A)^{-1} = 0 \tag{14}$$

用（$I-\rho A$）右乘（14）就得到了线性约束

$$e1'\ (I-\rho A)\ - e2'A = 0 \tag{15}$$

按照通常方式使用 Wald 检验就可以对这些约束条件进行估计。因此，如果理性定价公式成立且代表性投资者遵从理性预期假设，那么我们期望（14）~（15）的约束条件成立。（15）中的约束条件为：

$$(1,\ 0)\ =\ (\rho a_{11},\ \rho a_{12})\ +\ (a_{21},\ a_{22}) \tag{16}$$

也就是说，

$$1 = \rho a_{11} + a_{21} \tag{16a}$$

$$0 = \rho a_{12} + a_{22} \tag{16b}$$

单期收益不可预测

对于线性约束（16b），人们几乎找不到任何直觉，但可以很容易地证明预期单期实际超额收益 $E_t\ (h_{t+1}-r_{t+1})$ 不可预测，等价地说，市场不存在异常利润机会。使用（4）并忽略常数项就得到

$$\begin{aligned} E_t\ (h_{t+1}-r_{t+1} \mid \Omega_t)\ &= \delta_t - \rho E_t\delta_{t+1} + E_t\ (\Delta d_{t+1} - r_{t+1}) \\ &= \delta_t - \rho E_t\delta_{t+1} - E_t\ (rd_{t+1}) \end{aligned} \tag{17}$$

根据（11），

$$E_t\delta_{t+1} = a_{11}\delta_t + a_{12}rd_t$$

$$E_t\ (rd_{t+1})\ = a_{21}\delta_t + a_{22}rd_t$$

因此

$$E_t\ (h_{t+1}-r_t \mid \Omega_t)\ = \delta_t - \rho\ [a_{11}\delta_t + a_{12}rd_t]\ -\ (a_{21}\delta_t + a_{22}rd_t)$$

$$= (1-\rho a_{11}-a_{21})\ \delta_t - (a_{22}+\rho a_{12})\ rd_t \tag{18}$$

于是，在 Var 预测方程给定的条件下，根据 t 时刻所获得的信息，期望超额单期收益是可预测的，除非 Wald 检验中的线性约束（16）成立。对于非线性 Wald 检验的经济解释，我们将在这一章随后的部分中进行讨论。

“理论”红利价格比

我们将（13）的右边定义为理论价格红利比 δ'_t，使用无约束 Var 估计来获得 δ'_t 的时间序列

$$\delta'_t = \sum_{j=0}^{\infty} \rho^j\ (r^e_{t+j+1} - \Delta d^e_{t+j+1}) \ = e2'A\ (I-\rho A)^{-1} z_t \tag{19}$$

在理性定价公式和理性预期假设下，我们期望实际红利价格比 δ_t 能反映 δ'_t。我们从三个方面来对这个命题进行考察：（ⅰ）观察 δ_t 和 δ'_t 的时间序列图；（ⅱ）标准差比率 $SDR=\sigma\ (\delta'_t)\ /\ \sigma\ (\delta_t)$ 应等于 1；（ⅲ）相关系数 Corr（δ'_t，δ_t）应等于 1。我们不使用红利价格比而是使用恒等式 $p'_t=d_t-\delta'_t$ 来导出理论价格序列，将这个价格序列与由（ⅰ）~（ⅲ）所度量的实际股票价格进行比较。

Var 方法的深层含义

在 Var 方法假设的信息集合给定的条件下，我们所构造的变量 p'_t 体现了投资者对未来红利和时变收益（贴现率）DPV 的最佳预测。因此，它与 Shiller 最初的波动率检验中完全预见价格的期望值 $E_tP^*_t$ 非常接近。两者的差异在于，P^*_t 的计算不用特殊的 Var 方程来对基础变量进行估计，只需借助理性预期下的无偏性假设（$D_{t+1}=E_tD_{t+1}+w_{t+1}$）。换句话说，$p'_t$ 以特定的红利统计模型为条件，而 $E_tP^*_t$ 却不是。

简要分析（对数）完全预见价格 $p^*=\ln\ (P^*_t)\ t$ 和（对数）理论价格 p'_t 之间的关系是值得的，我们至少可以清楚地知道两个相关概念之间的区别。因此，我们可以看出 Var 方法相对于 Shiller 方差边界检验和基于 P^*_t 回归检验的优缺点。对数线性恒等式（2）（用 r_{t+1} 代替 h_{t+1}）可以重新排列为

$$p_t = \rho p_{t+1} + \ (1-\rho)\ d_{t+1} - r_{t+1} + k \tag{20}$$

将方程（20）向前递归迭代就可以得到用实际未来红利和贴现率所表示的 DPV 对数线性表达式

$$p^*_t = \ (1-\rho) \sum_{j=0}^{\infty} \rho^j d_{t+j+1} - \sum_{j=0}^{\infty} \rho^j r_{t+j+1} + k/\ (1-\rho) \tag{21}$$

方程（21）使用的是实际（事后）值，因此，p^*_t 代表的是（对数）完全预见价格。有效市场－理性定价公式（EMH－RVF）假设意味着：

$$p_t = E\ (p^*_t \mid \Omega_t) \tag{22}$$

在理性预期下，代表性投资者使用所有可用信息 Ω_t 来计算 $E_tp^*_t$，但理论价格 p'_t 的计算只使用了包含在 Var 中的受限信息集合，受限信息集合由计量经济学家来进行选择。“现实”中的投资者使用比 Var 更多的信息，但这是会使 Var 的结果“弱于”基于完全预见价格 p^*_t 和 Shiller 方差不等式的检验吗？在某些情形中确实如此，但在其他情形中则不尽其然。

首先，Var 包含 δ_t，这也是在 Var 方法中 Shiller 方差边界不等式 Var（P_t）≤Var（P^*_t）可以转化等式 $p_t=p'_t$ 的原因。此外，即使在 Var 中增加更多变量，我们仍然期望系数 δ_t 等于 1。为了看清这一点，注意，由（19）有

$$\delta'_t = e2'f\ (a)\ z_t \tag{23}$$

$$z_t = \ [\delta_t,\ rd_t]$$

$$f\ (a)\ = A\ (I-\rho A)^{-1} \tag{24}$$

对滞后一阶的 Var 而言，$f(a)$ 是一个（2×2）矩阵，是 Var 系数 a_{ij} 的非线性函数。$f(a)$ 的第二行是 2×1 向量 $[f_{21}(a), f_{22}(a)] = e2'f(a)$，$f_{21}$ 和 f_{22} 是参数 a_{ij} 的标量（非线性）函数。于是，由（23）我们有

$$\delta_t' = f_{21}(a)\ \delta_t + f_{22}(a)\ rd_t \tag{25}$$

因为在有效市场假设下，$\delta_t = \delta_t'$，我们期望标量系数 $f_{21}(a) = 1$，$f_{22}(a) = 0$。即使在 Var 中增加变量，这些约束条件仍然成立。如果在 Var 系统中增加一个变量，那么方程顶多就是增加一项 $f_3(a)y_t$，理性定价公式意味着 $f_3(a) = 0$（除上述两个约束以外）。

因此，如果 Var 限制条件在受限信息集合中遭到拒绝，那么在 Var 的更大信息集合中限制条件仍然会遭到拒绝。就逻辑而言，后者是很自然的情形，如果样本数据足够的多，那么这种情形就是真实情形。在这种意义下，受限信息集合的使用并不是主要的缺陷。然而，在有限数据集合中，我们知道计量经济学家有可能错误地从回归方程中删掉一些变量，代表性投资者在进行预测时也可能犯这样的错误，由此会导致结果是“不正确”（即不一致）的参数估计。因此，在 Var 中使用更多变量有助于 RVF 有效性检验效果的提高。

因此，基于完全预见价格的 Shiller 方差边界检验可能存在一些问题，因为数据存在非平稳性问题，基于 Var 方法的结果也可能因为省略了一些变量或存在一些错误设定（如错误的函数形式）而导致了一些问题。尽管我们可以用诊断性检验（如对 Var 误差项的序列相关性进行检验等）来考察 Var 表示的统计有效性，但这种方法在有限样本中会导致错误的推断。

我们现在试图获得一些直觉来对（14）中非线性约束的 Wald 检验进行经济解释，对双变量 Var 而言，（14）中的约束等价于（25）中的约束。如果非线性约束遭到拒绝，那么可能的原因是 $f_{22}(a) \neq 0$。如果真是这样，那么 rd_t 就会对 $\sum_{j=0}^{\infty}\rho^j(h_{t+j+1}^e - \Delta d_{t+j+1}^e)$ 产生影响。从某种意义上说，单期收益加权是多期收益的一种形式，对线性约束的违反意味着长期（加权）收益可预测。我们在此不希望对这一点进行深入讨论，因为这涉及下面章节中的多期收益。

我们可以用 Var 方法来给出对有效市场－理性定价公式（EMH－RVF）假设有效性进行评估的另外一种方法。这种方法的基础是，将 $r_t - \Delta d_t$ 进行分解使得 $z_t = (\delta_t, \Delta d_t, r_t)'$。然后将（19）的右边进行分解

$$\delta_t = \delta_{dt}' + \delta_{rt}' \tag{26}$$

其中，

$$\delta_{rt}' = e3'A\ (I - \rho A)^{-1}z_t \tag{27a}$$

$$\delta_{dt}' = e2'A\ (I - \rho A)^{-1}z_t \tag{27b}$$

因此，我们预期相关系数 $\mathrm{Corr}(\delta_t - \delta_{dt}', \delta_{rt}') = 1$。如果这个相关系数大大小于 1，那么这意味着实际预期收益 δ_{rt}' 不能充分解释红利价格比的运动（对未来红利预测 $\delta_t - \delta_{dt}'$ 的影响要进行纠正）。

正如我们即将看到的那样，基于（26）的方差分解可以用来考察预期收益的持续性对红利价格比 δ_t 和股票价格（$p_t = d_t - \delta_t$）的影响。预期收益持续性的程度可以用 Var 矩阵 A 中某些系数的大小来进行建模。我们可以用（26）将 δ_t 的变化分解为

$$\mathrm{Var}(\delta_t) = \mathrm{Var}(\delta_{dt}') + \mathrm{Var}(\delta_{rt}') + 2\mathrm{Cov}(\delta_{dt}', \delta_{dt}') \tag{28}$$

右边的项是 Var 矩阵 A 的函数。然而，我们在这里不展开分析，协方差项在这一节中不会出现，因为我们的目的是将 δ_{rt} 和 $\delta_t - \delta_{dt}'$ 进行比较。

小结：理性定价公式与向量自回归

我们在前面讲了一大堆，Var 方法在股票价格和收益中的应用要点如下：

（ⅰ）线性化理性定价公式（RVF）意味着 Var 参数存在一组交叉方程限制，这些交叉方程限制确保了单期超额收益不可预测，理性预期下的预测误差独立于 t 时刻的信息。

（ⅱ）使用 Var 方法可以计算出“理论”红利价格比 δ_t' 和理论股票价格 $p_t' = (d_t - \delta_t')$。在

有效市场－理性定价公式（EMH－RVF）假设下，我们期望 $\delta_t=\delta'_t$，$p_t=p'_t$。可以用图来对它们进行比较。

（iii）标准差比率 $SDR=\sigma(\delta_t)/\sigma(\delta'_t)$ 和相关系数 (δ_t, δ'_t) 都应等于1。δ_t 是 $r_t-\Delta d_t$ 未来变化的充分统计量，因此，δ_t 至少是后者的 Granger 原因。

（iv）Var 可以用来度量预期红利（δ'_{dt}）和预期未来收益 δ'_{rt} 对红利价格比 δ'_t 变化的相对重要性。

12.2　实证结果

我们陈述的所有结果都是为了说明问题。因此如果证据不是很充足，那么这些结果就不是肯定成立的结果。实证研究主要讨论以下问题：

（i）预期的单期持有期收益模型的选择问题；

（ii）Var 模型中应包含多少个变量，合适的滞后长度为多大，参数估计是否具有短暂稳定性（temporal stability）；

（iii）如何解释各种“度量”方法之间相互矛盾的结果，如单一方程、相关性、方差比统计量和 Var 方法的 *Wald* 检验等研究中单期收益的可预测性问题。

理性定价公式与收益可预测性

Campbell 和 Shiller（1988）的学术研究使用的是1871—1986年期间美国股票指数和相关红利的年度数据。对单期必要收益而言，他们使用了四种不同的假设：

（a）必要实际收益为常数（即 h_t 为常数）；

（b）必要名义（或实际）收益等于名义（或实际）国库券利率，$h_t=r^c_t$；

（c）必要实际收益与消费增长率成比例（即 $h_t=\gamma\Delta c_t$，γ 为相对风险规避系数）；

（d）必要实际（市场）收益取决于风险溢价，风险溢价等于相对风险规避系数乘以“市场投资组合”收益的（条件）方差 V_t：

$$h_t=\gamma V_t$$

因此，在 Var 中 h_t 可以用上述假设中的任何一个来进行替换。注意，（b）是无风险溢价的通常性假设，（c）的基础是基于消费的资本资产定价模型（CCAPM），松散地说（d）中风险溢价的基础是以为市场投资组合为参照系的 CAPM 模型（V_t 是事后收益的方差，作为风险的一种度量，它是市场收益条件方差一个非常粗糙的度量指标）。

不管我们选择（a）～（d）中的哪一个必要收益假设，结果都不会有本质的变化，因此我们主要讨论假设（a），即实际收益为常数（Campbell 和 Shiller（1988），表（4））。变量 δ_t、r_t 和 Δd_t 是平稳的变量 $I(0)$。在 δ_t 和 Δd_t 的信息集合上，（近似）对数收益的单方程回归为：

$$h_t=0.141\delta_t-0.012\Delta d_t \tag{29}$$

$$(0.057)\quad(0.12)$$

$$1871—1986,\ R^2=0.053\ (5.3\%),\ (.)\ =\text{标准差}$$

在解释年（单期）实际收益时，只有红利价格比在统计上显著，但解释力很弱（$R^2=5.3\%$）。在使用 $z_{t+1}=(\delta_{t+1}, \Delta d_{t+1})$ 的 Var（滞后长度为1）中，变量 δ_{t+1} 具有很强的自回归性质，解释力（$R^2=0.515$）主要来自于 δ_t（系数等于0.706，标准误差为0.066），而 Δd_t 对解释基本上没有贡献。实际红利的变化 Δd_{t+1} 对 Δd_t 有部分解释力，但红利价格比 δ_t 在统计上显著，且有“正确的”符号（见（5））。因此，δ_t 是 Δd_t 的 Granger 原因，这是 RVF 的一种弱检验。如果使用估计的 Var 矩阵 A，并用（23）来计算（25）中的 $f_{21}(a)$ 和 $f_{22}(a)$，那么 Campbell－Shiller 所发现的结果为

$$\delta'_t=0.636\delta_t-0.097\Delta d_t \tag{30}$$

$$(0.123)\quad(0.109)$$

在有效市场-理性定价公式（EMH-RVF）假设下，我们期望δ_t的系数为1，Δd_t的系数为0：尽管后者通过了检验，但前者却遭到了拒绝。根据理论分析，如果$\delta_t' \neq \delta_t$，那么单期收益具有可预测性，Var结果与单方程回归结果一致（单方程回归的结果是单期收益具有可预测性（如Fama-French（1988b）））。根据标准差比率为1的假设，我们要拒绝交叉方程的Wald检验：

$$\mathrm{SDR} = \sigma(\delta_t')/\sigma(\delta_t) = 0.637 \quad (s.\ e.\ = 0.12) \tag{31}$$

然而，δ_t和δ_t'之间的相关性高达0.997（*s. e.* = 0.006），这与1在统计上存在显著差异。因此，δ_t和δ_t'的运动方向相同，但实际δ_t的变化比理论预测值δ_t'（在理性定价公式下）约大60%（即1/0.637 = 1.57）。因此，即使我们考虑红利和贴现率的时变性，红利价格比和股票价格的波动性仍然比“基础变量”大。

如果使用单方程回归（29），那么我们可以在4.5%的显著性水平下拒绝零假设“单期收益不可预测”，然而，如果使用Var方法，那么我们将在一个更高的显著性水平0.5%下拒绝$\delta_t = \delta_t'$的Wald检验。注意，δ_t'是所有未来单期收益h_{t+1}的加权平均，也是长期收益的一种逼近。因此，拒绝Wald检验的结果Fama-French的结果一致（Fama-French的研究结果是长期收益比短期收益具有更强的可预测性）。

研究结果也表明，尽管单期收益几乎不可预测，这还是意味着等式$\delta_t = \delta_t'$（或$p_t = p_t'$）大体上遭到了违反。因此，尽管单期收益在很大程度上不可预测，实际股票价格的波动性还是比由理性定价公式所预测的股票价格的波动性高。

对不同的预期收益模型而言，当Var模型的滞后阶数大于1时，$(\delta_t - \delta_{dt}')$和$\delta_{rt}'$之间的相关性基本上都比较低。因此，一个试探性的结论是，预期未来收益的波动率无法充分解释实际股票价格的波动。在这种模型变体中，股票价格的变化在很大程度上要归因于预期红利的变化，尽管后者无法“完全”解释股票价格的波动（即$\mathrm{Var}(\delta_t) > \mathrm{Var}(\delta_t')$且$\mathrm{Var}(p_t) > \mathrm{Var}(p_t')$）。

在后来的研究中，Campbell和Shiller（1988）（第8章）将Var的信息集合进行扩大，使之包含（对数）收益价格比e_t，其中$e_t = \bar{e}_t - p_t$，$\bar{e}_t$等于实际收益对数值的长期移动平均。将e_t包含进来的动机是金融分析师经常使用PE（市盈率，或盈利收益率*E/P*）来预测股票收益。事实上，Campbell-Shiller所发现的结果是，盈利收益率是决定收益h_{t+1}的关键变量，从统计角度来看，它的表现优于红利价格比。现在的Var包括三个“基础”变量$z_{t+1} = (\delta_{t+1}, \Delta d_{t+1}, e_{t+1})'$，但Var的基本结果基本上没有改变。

自20世纪50年代后期开始，股票的实际价格和理论价格之间存在明显差异，因此我们应拒绝理性定价公式。股票价格存在过度波动性，尽管实际单期（对数）收益h_t与理论收益h_t'高度相关（例如，Corr$(h_t, h_t') = 0.915$，*s. e.* = 0.064，参见Shiller（1988）（表2）），股票价格与“基础价值”p_t'之间仍然有很大的差异。使用（19）和$p_t' = d_t - \delta_t'$来计算理论价格p_t'就可以看出上述结果的原因。

$$p_t' = 0.256p_t + 0.776e_t + 0.046d_t - 0.078d_{t-1} \tag{32}$$

因此，在p_t'的决定式中，p_t的权重仅为0.256而不是1，p_t'的长期变动受盈利e_t的“匀滑”移动平均所主导。然而，就短期而言，p_t的波动率很高，这使得p_t'的波动率也很高。根据定义，单期收益在很大程度上取决于价格变化，h_t和h_t'高度相关。

无论必要（实际）收益或相对于商业票据利率的超额收益是否用作时变贴现率，Campbell-Shiller的结果基本上没有变化。就1927—1986年数据集合的各个子集而言，不同滞后长度的Var模型不会改变上述定性结果。然而，蒙特卡罗模拟表明（Campbell和Shiller（1989），Shiller和Beltratti（1992）），当滞后长度“太大”（大于3）时，Wald检验会频繁地拒绝RVF成立的零假设。尽管蒙特卡罗模拟得出了这样的结论，Campbell和Shiller（1989）还是注意到，在1000次模拟中，Wald检验没有遭到一次拒绝，这与实际数据集合中(δ_t, δ_t')、(p_t, p_t')或(h_t, h_t')之间相关性很低时的结果差不多。这意味着，尽管Var方法存在一定的偏差，

实际数据违背假设的结果比“EMH - RVF 是正确零假设下”人们所期望的结果更糟糕。

Cuthbertson、Hayes 和 Nitzsche（1997）使用英国股票价格指数的年度数据（1918—1993）所发现的结果与 Campell 和 Shiller（1988）常数超额收益模型的结果类似。然而，CAPM 模型意味着超额市场收益取决于市场波动率，即 $E_t h_{t+1} - r_t = \gamma E_t V_{t+1}$。波动率用市场收益的方差来进行度量，可以通过最小化 Wald 统计量来确定相对风险规避系数（这里选择的 $\gamma = 2.5$，但 γ 通常在 1 到 10 之间取值）。他们发现波动率 V_t 是 Δd_{t+1} 的 Granger 原因，它部分地决定了理论红利价格比 δ_t'。因此，时变波动率 V_t 传递了 δ_t' 更多的波动率，相对于常数收益模型而言，这个模型与实际 δ_t 的变动更加一致。

表 1　　常数预期超额收益（英国 1918—1993 年期间的年度数据）

	收益时期（年）					
	1	2	3	5	10	∞
（ⅰ）Wald 统计量	16.137	24.722	33.227	43.281	48.484	48.725
（p - 值）	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
（ⅱ）$\sigma(\delta_t')/\sigma(\delta_t)$	0.645	0.447	0.353	0.291	0.275	0.275
（$s.e.$）	(0.144)	(0.150)	(0.146)	(0.138)	(0.133)	(0.133)
（ⅲ）Corr (δ_t, δ_t')	0.972	0.959	0.961	0.975	0.981	0.981
（$s.e.$）	(0.020)	(0.049)	(0.052)	(0.033)	(0.021)	(0.021)
（ⅳ）Corr (h_{1t}, h_{1t}')	0.963	0.929	0.899	0.866	0.852	0.852
（$s.e.$）	(0.027)	(0.048)	(0.076)	(0.106)	(0.113)	(0.113)

注：

1. 表中的统计量由三变量 Var 模型导出，三个变量分别是 δ_t、r_t 和 Δd_t。
2. 原始数据来自于 Barclays - deZoete - Wedd（BZW）所著的 Equity Gilt Study，在 Barclays Capital 中可以找到这些数据。
3. 标准差是对异方差进行修正后的标准差。

对常数收益模型（表 1）而言，比率 $\sigma(\delta_t')/\sigma(\delta_t)$ 小于理论值 1，Wald 检验在所有预测期都遭到了拒绝（见下）。然而，如果将波动率加入到 Var 模型中，那么 $\sigma(\delta_t')/\sigma(\delta_t)$ 的点估计在所有预测期上都大于 1，但又与 1 在统计上无显著差异（因为标准误差太大）。此外，Wald 检验也没有遭到了拒绝（表 2）。

CAPM 模型的波动率项确实增加了 Var 的解释力，V_t 在 Δd_{t+1} 和 δ_{t+1} 的 Var 方程中是统计显著的。将图 1 和图 2 进行比较就可以看出，如果模型包含波动率，那么 δ_t' 和 δ_t 之间的关系会出现显著变化。

表 2　　CAPM 波动率模型（$\gamma=2.5$）

	收益时期（年）					
	1	2	3	5	10	∞
（ⅰ）Wald 统计量	2.698	2.118	1.867	1.728	1.708	1.711
（p－值）	(0.441)	(0.548)	(0.600)	(0.631)	(0.635)	(0.635)
（ⅱ）$\sigma(\delta'_t)/\sigma(\delta_t)$	1.350	1.666	1.844	2.007	2.091	2.098
（s.e.）	(0.331)	(0.623)	(0.803)	(0.977)	(1.067)	(1.076)
（ⅲ）Corr（δ_t，δ'_t）	0.945	0.945	0.945	0.946	0.946	0.946
（s.e.）	(0.053)	(0.054)	(0.054)	(0.055)	(0.055)	(0.055)
（ⅳ）Corr（h_{1t}，h'_{1t}）	0.898	0.884	0.879	0.875	0.874	0.873
（s.e.）	(0.127)	(0.135)	(0.138)	(0.143)	(0.144)	(0.145)

注：

1. “CAPM” 模型的超额收益取决于市场波动率，即 $E_t h_{t+1} - r_t = \gamma E_t V_{t+1}$，波动率 V_{t+1} 用股票年收益的方差来进行度量。

2. Var 包含变量 $\{\delta_t, \Delta d_t, V_t\}$。

3. 标准差是对异方差进行修正后的标准差。

使用 CAPM 波动率模型（股票指数，1973.6～1996.6 的月度数据）对法国和德国进行研究（Cuthbertson 和 Hyde（2002））的结果比常数超额收益模型好。就法国而言（图 3a），波动率模型支持理性定价公式；就德国而言，结果就要差一些（图 3b）。Cuthbertson 和 Hyde 认为原因是德国市场的波动性比法国弱，德国市场的波动率不足以对预期收益产生影响（见第 29 章关于 GARCH－M 模型的讨论）。

Cuthbertson、Hyde 和 Nitzsche（1997）注意到，在使用 Var 时“各种假设”都具有普遍性（即市场无效率）。因此，他们用 Var 方法来对效率的一种特殊形式（即英国股票市场上的短视主义）进行检验。理性定价公式可以写为

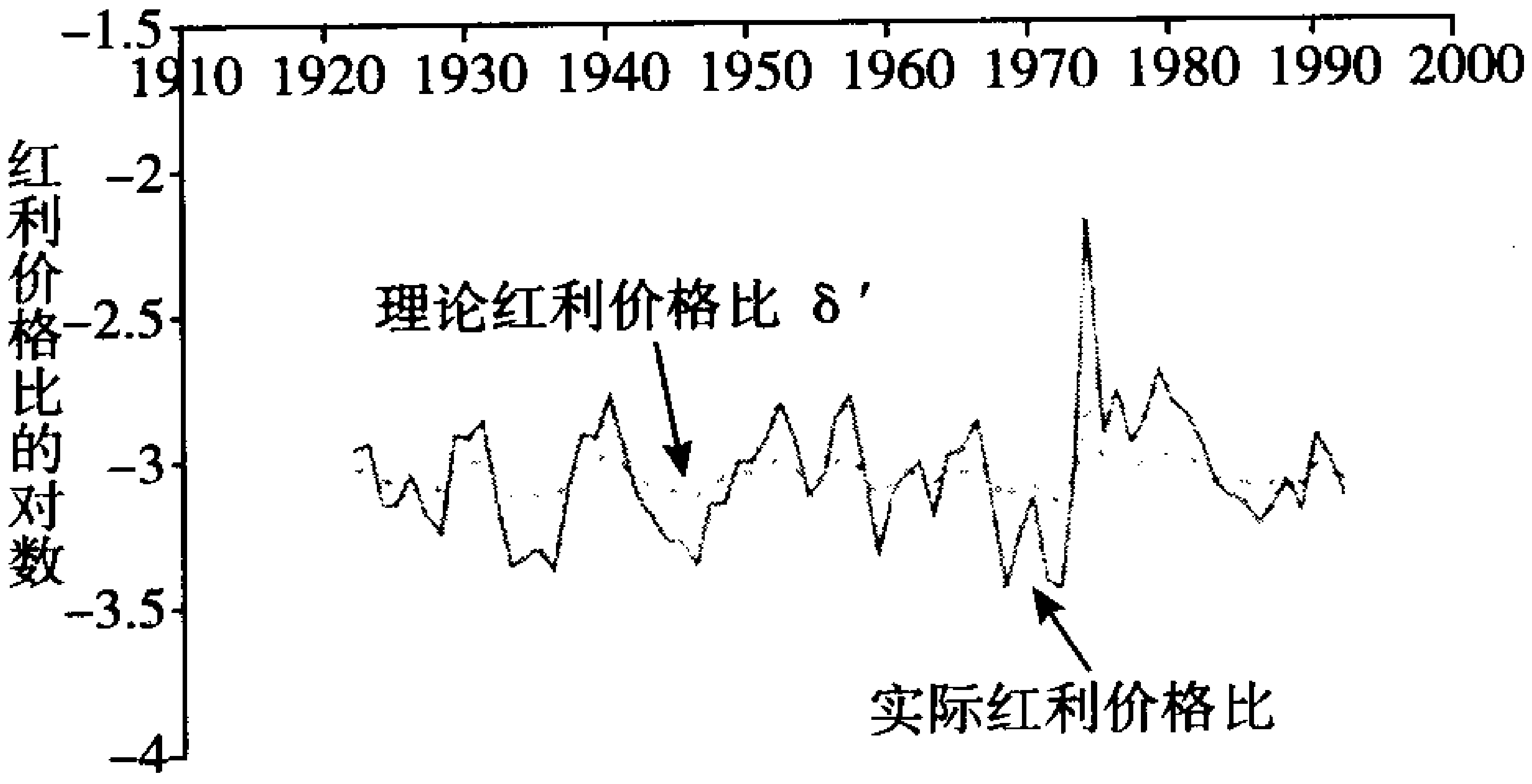


图 1　实际和“理论”红利价格比——常数收益模型（英国数据：1921—1993）

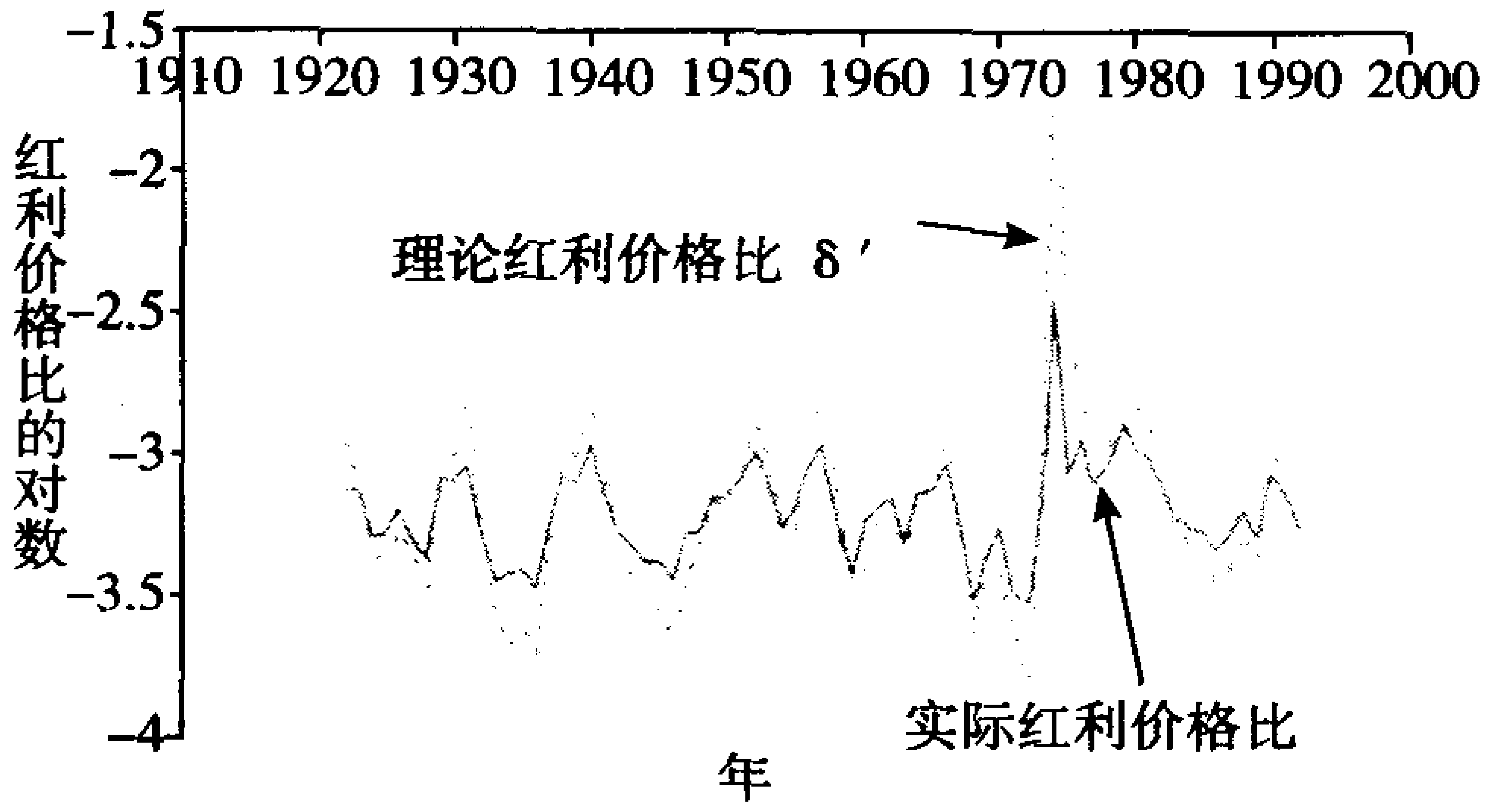
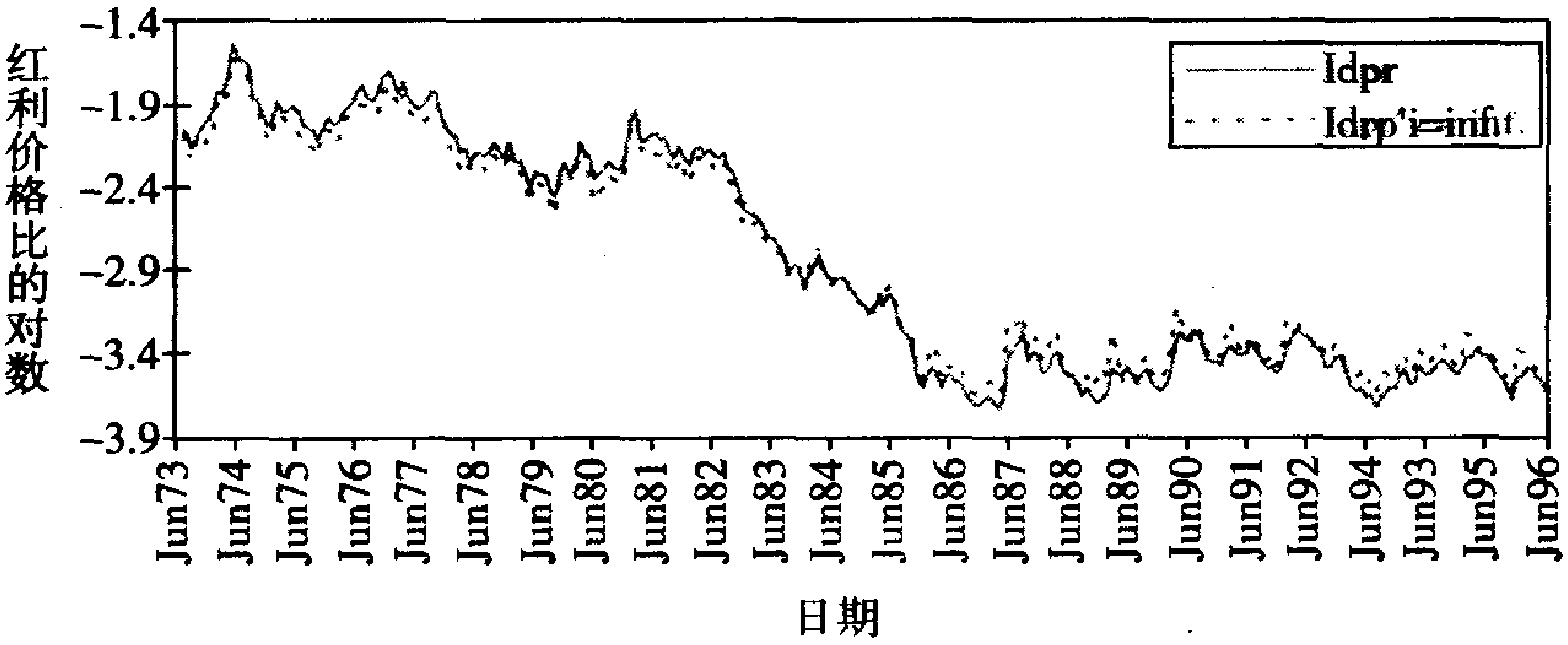
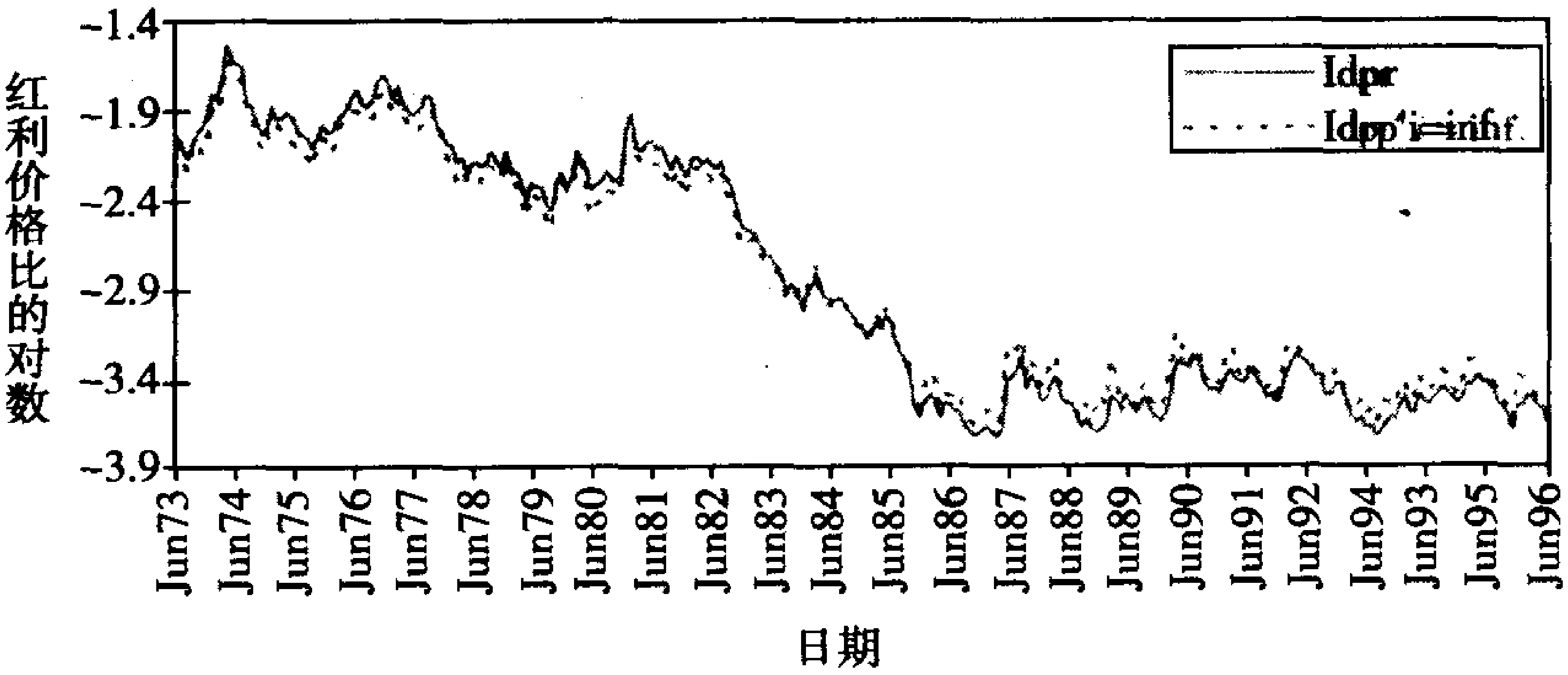


图 2　实际和"理论"红利价格比——CAPM 波动率模型，
$\gamma = 2.5$（英国数据：1921—1993）



a) 实际和理论红利价格比对数值的比较

（市场波动率 -CAPM 模型：法国）



(b) 实际和理论红利价格比对数值的比较

（市场波动率 -CAPM 模型：德国）

图 3

注：

（ⅰ）Idpr 等于实际红利价格比的对数，Idpr（i = inf）是 RVF 中预测期为无穷大时的理论红利价格比的对数；

（ⅱ）"市场波动率 - CAPM 模型"假设预期超额收益取决于市场收益的条件波动率（后者用历史收益的方差来进行度量）；

（ⅲ）法国和德国情形中所使用的相对风险规避系数为 $\gamma = 4$。

$$\delta_t = \sum_{j=0}^{4} x^{j+1} E_t Y_{t+1+j} + x^6 \rho^5 E_t \delta_{t+5}$$

其中，$Y_{t+1+j} \equiv \rho^j\ (h_{t+1+j} - \Delta d_{t+1+j})$。市场有效要求 $x=1$，短视主义（在 EMH 下，未来“基础变量”的权重更小一些）意味着 $x<1$。因为 Var 方法允许出现时间序列 Y_{t+1+j}和 δ_{t+j}的预测变量，所以我们可以用上述方程来估计 x。注意，我们可以对 δ_{t+j}的预测值进行截断，比如说在 $j=5$ 年处进行截断，或者可以用

$$E_t \rho^5 \delta_{t+5} \equiv E_t \sum_{j=5}^{\infty} \rho^j\ (h_{t+1+j} - \Delta d_{t+1+j})$$

作为 $E_t \rho^5 \delta_{t+5}$的度量。使用 CAPM 波动率模型的 Var 预测（$\gamma = 2.5$）所得到的 x 的估计值约为 0.7（*s. e.* =0.7），这意味着短视行为确实存在。因此，当我们假设超额市场收益依赖于市场波动率时，尽管理性定价公式并没有遭到拒绝，但短视行为的存在却拒绝了有效市场假设。

Cuthbertson、Hayes 和 Nitzsche（1999）使用 Var 方法对英国股票总收益的季度数据（1965—1992）进行了检验，结果明确地拒绝了 CCAPM 模型（即 $E_t h_{t+1} - r_t = \gamma + E_t \Delta c_{t+1}$），但 $\gamma > 3$ 的 CAPM 波动率模型却表现不错。他们接下来对五个市场板块（如资本品、消费品等）进行了考察，根据 CAPM 模型，任何板块 i 的预期收益都应取决于资产 i 收益与市场收益之间的协方差：

$$E_t h_{i,t+1} - r_t = \gamma \mathrm{Cov}_t\ (h_{i,t+1},\ h_{m,t+1})$$

时变条件协方差由交叉乘积 $(h_i h_m)_{t+1}$来度量，Cuthbertson 和 Galindo（1999）将 Var 应用于板块收益。当我们假设板块收益取决于板块波动率而不是各板块收益与市场收益之间的协方差时，令人感到惊讶的是，使用 Var 度量的各板块结果得到了极大的改善，这意味着投资者将各个板块都看作市场的一个组成部分。这可能是因为，“营造市场的”投资者实际上只关注一个板块（即板块专家），所以某一板块的波动率对他们来说更为重要，这也是板块收益的影响因素（即市场分割）之一。

正如我们在第 18 章中即将看到的那样，上述结果也可以用行为金融学的理论解释。研究者（在实验情况中）观察到，人们将单个赌博的结果与其他财富分开对待（假设没有制度约束）。例如，当面临一项赌博时，人们将它视为“独立的赌博”，而不是将它与其他“赌博”合在一起进行考虑。这是“心理账户”（mental accounting）的一个例子，也就是众所周知的狭窄框架（narrow framing）。在上述例子中，投资者应将在所有板块中的赌博合在一起进行考虑，他们应该意识到任意板块的“赌博”风险都依赖于这个板块的收益与市场收益之间的协方差。但是，如果他们属于“狭窄框架”类型，那么他们将任意板块中的赌博风险分别对待。因此，板块收益的方差是投资者感知风险的正确度量。

一个非常有趣的个体数据研究是 Bulkley 和 Taylor（1996），他们用 Var 来进行预测（即理论价格以一种有趣的形式出现）。首先，用 1960—1980 年的数据递归地估计每家公司 i 的理论价格 P'_{it}，对每一年的递归样本而言，理论价格 P'_{it}与实际价格 P_{it}之差有助于预测公司一年到十年期间的收益（要对一些公司的风险变量 z_k 进行修正）：

$$R_i = \alpha + \gamma_\theta\ (P'_i / P_i)\ + \sum_{k=1}^{m} \gamma_k z_k$$

与有效市场假设相反，他们发现 $\gamma_0 \neq 0$。另外，他们将公司根据（P'_i / P_i）来进行排序，选出最好的/最差的 20 家公司（或最好的/最差的 10 家公司）来构造投资组合。持有由 20 家最好公司构成的投资组合三年所获得的超额收益比 S&P500 指数的收益每年约高 7%。他们也发现，20 家最好公司的股票价格出现了约 25% 的定价偏差。

Shiller 波动率检验与多期收益

在对 Shiller 波动率不等式和 Fama - French 长期回归的讨论中，我们注意到 $P_t - P_t^*$（这里的 P_t^* 是完全预见价格）与“长期”收益类似。因此，如果长期收益不可预测，那么 $P_t - P_t^*$

对信息 Ω_t 的回归系数应为零。Fama - French 使用 N 期实际长期收益 R^N_{t+N}发现，历史收益（特别是 3 ~ 5 年期收益）有助于它的预测。Fama - French 在他们的检验中使用的是单变量 AR 模型。

Campbell 和 Shiller（1988）（第 8 章）将单期收益的线性方程应用到多期收益上，这意味着 Var 系数存在交叉方程限制。因此，使用 Var 方法，人们可以在多变量框架中考察 Fama - French 的“长期”结果。从 t 到 $t+i$ 的第 i 期收益可以表示为

$$(1+H_{i,t}) \equiv (1+H_{1,t+1})(1+H_{1,t+2})(1+H_{1,t+3})\cdots(1+H_{1,t+i})$$

其中，$H_{1,t+j}$是从 $t+j-1$ 到 $t+j$ 的单期收益。因此

$$h^*_{i,t} = \sum_{j=0}^{i-1} h_{1,t+j+1} \tag{33}$$

其中，$h^*_{i,t} = \ln(1+H_{i,t})$。随着期限 i 的增加，方程（33）变成无穷大，因此 Campbell - Shiller 倾向于选择 i 期对数收益的加权平均值进行讨论

$$h_{i,t} = \sum_{j=0}^{i-1} \rho^j h_{1,t+j+1} \tag{34}$$

对单期收益 $h_{1,t+1}$而言，使用方程（34）和恒等式（4）就得到

$$\begin{aligned} h_{i,t} &= \sum_{j=0}^{i-1} \rho^j (\delta_{t+j} - \rho\delta_{t+1+j} - \Delta d_{t+1+j} + k) \\ &= \delta_t - \rho^i \delta_{t+i} + \sum_{j=0}^{i-1} \rho^j \Delta d_{t+1+j} + k(1-\rho^i)/(1-\rho) \end{aligned} \tag{35}$$

方程（35）是关于 δ_t、δ_{t+1}和 Δd_{t+j}的（近似）恒等式，它定义了从 t 到 $t+i$ 的多期收益 $h_{i,t}$。这似乎没有太多的直观含义，但我们简要地解释前三项（最后一项是线性化常数，没有实质性意义）。单期收益取决于当前的红利价格比，这可以解释（35）中的第一项；多期收益必须依赖于从 t 到 $t+i$ 的红利支付，如果我们通过 δ_t（$=d_t-p_t$）来确定 d_t，那么它必须取决于红利增长率，这就是（35）右边的第三项；随着收益期的增加，第二项的重要性逐渐下降（因为，如果 δ 是平稳序列且 $0<\rho<1$，那么 $\rho^i\delta_{t+i}\to 0$），它之所以出现是因为有限时期上的收益取决于 $t+i$ 时刻的终点（卖出）价格，也取决于 δ_{t+i}。

我们现在来看看如何将基于 Var 的多变量预测模型与 Fama - French 的“长期”单方程回归模型进行比较。包含 $z_t=(\delta_t, \Delta d_t)$ 的 Var 可以用来预测（35）右边第 i 期的理论收益，记为 $h'_{i,t}$。我们可以用图、方差比率及 $h'_{i,t}$与 $h_{i,t}$之间的相关性来对 $h'_{i,t}$和 i 期收益 $h_{i,t}$进行比较。

尽管在包含 δ_t 和 Δd_t 的 Var 模型基础上可以用（35）来对 $h_{i,t}$进行预测，方程（35）并没有给出对 A 矩阵进行约束的 Wald 检验，因为 $h_{i,t}$不属于信息集合。然而，稍微修改一些就可以得到多期收益的 Wald 检验。我们引入预期单期超额收益为常数的行为假设

$$E_t(h_{1,t} - r_t) = c \tag{36}$$

于是，

$$E_t(h_{i,t} \mid \Omega_t) = \sum_{j=0}^{i-1} \rho^i E_t r_{t+j+1} + c(1-\rho^i)/(1-\rho) \tag{37}$$

对（35）求数学期望，令（35）的右边等于（37）的右边，我们就得到熟悉的 δ_t 的差分方程，向前求解就可以得到第 i 期的红利价格比

$$\delta_t = \sum_{j=0}^{i-1} \rho^j E_t(r_{t+1+j} - \Delta d_{t+1+j}) + \rho^i E_t \delta_{t+i} + (c-k)(1-\rho^i)/(1-\rho) \tag{38}$$

如果我们忽略常数项，那么（38）与前面得到的表达式类似，不同之处在于，求和式子上标的最大值为 i 而不是无穷大，这就是 i 期中的动态 Gordon 模型。Campell 和 Shiller（1988）使用（38）针对不同的来构造多期收益的 Wald 检验。对 $z_t=(\delta_t, rd_t, e_t)'$而言，约束条件为

$$\mathrm{e1}'(\mathrm{I}-\rho^i \mathrm{A}^i) = \mathrm{e2}'\mathrm{A}(\mathrm{I}-\rho\mathrm{A})^{-1}(\mathrm{I}-\rho^i A^i) \tag{39}$$

当 $i=1$（或 $i=\infty$）时，上式变成

$$\mathrm{e1}'(\mathrm{I}-\rho\mathrm{A})' = \mathrm{e2}'\mathrm{A} \tag{40}$$

这就是前面详细讨论的情形。如果（40）成立，那么右乘（$I-\rho^iA^i$）后就可以看到（39）对任意 i 在都成立。这意味着，如果单期收益不可预测，那么 i 期收益也不可预测。

Campbell 和 Shiller 使用 1871—1987 年的 S&P 指数数据发现，$i=1$ 时的 Wald 检验只在 2% ~ 4% 的显著性水平下遭到拒绝，但在 $i>3$ 时遭到拒绝的证据却要强得多（即显著性水平低于 0.2%）。这就是 Fama - French 单方程检验结果（多期收益比单期收益的可预测性更强）的反映。Cuthbertson、Hayes 和 Nitzsche（1999）使用英国数据所发现的结果与常数实际超额收益情形中的结果相似（见表 1）。但是，正如上面所注意到的那样，他们也发现，如果（与 CAPM 模型中的情形一样）波动率对单期收益存在影响，那么多期收益不可预测（见表 2 中 Wald 检验的结果）。

完全预见价格与多期收益

我们现在来证明，如果多期收益 $h_{i,t+1}$ 可预测，那么 Shiller 方差边界不等式很有可能遭到违反。根据我们在（39）中对（加权）多期收益 $h_{i,t+1}$ 的定义，当 $i\to\infty$ 时，$h_{i,t+1}$ 仍然是有限数

$$\begin{aligned}\lim_{i\to\infty}h_{i,t+1}&=(1-\rho)\sum_{j=0}^{\infty}\rho^j d_{t+1+j}-p_t+k/(1-\rho)\\&=\ln P_t^{**}-\ln P_t+k/(1-\rho)\end{aligned}\tag{41}$$

我们将（41）右边第一项写成 $\ln P_t^{**}$，因为它与 Shiller 波动率检验中完全预见价格 P_t^* 的对数值等价。如果第 i 期收益 $h_{i,t+1}$ 给予 t 时刻信息集合 Ω_t 是可预测的，那么由（41）知，在（$\ln P_t^{**}-\ln P_t$）对 Ω_t 的回归分析中我们也能看到 Ω_t 在统计上是显著的。因此，$\ln P_t^{**}\neq\ln P_t+\varepsilon$，Shiller 方差边界不等式不成立。当 i 有限时，$h_{i,t+1}$ 所对应的方程（35）是 $\ln P_{it}^*-\ln P_{it}$ 的对数线性表达式，P_{it}^* 的计算假设 $t+i$ 时刻的终点完全预见价格等于实际价格 P_{t+i}。变量 $\ln P_{it}^*$ 与 Mankiw、Romer 和 Shapiro（1991）在进行波动率不等式检验时所用到的变量类似，他们在不同的投资期上计算完全预见价格。因此，对有限时间而言，从宽泛意义上讲，对 h_{it} 的检验大体上等价于 Mankiw 等人的波动率不等式检验。两项研究所给出的结果类似，但 Campbell 和 Shiller（1988）对 EMH 的拒绝程度比 Mankiw 等强。

对股票价格水平而言波动率检验遭到违反的结论（根据 Shiller（1981）的早期工作）——也就是说 Var（P_t）>Var（P_t^*）——与长期收益可预测的观点是一致的。对（39）和（40）的多期情形而言，Var 中 Wald 约束检验遭到拒绝的结果基本上与 $j\to\infty$ 时 Shiller 方差边界检验遭到违反的结果是一致的，这也是对多期收益不可预测假设的一种反驳（与 Fama 和 French（1988b）及 Mankiw、Romer 和 Shapiro（1991）的发现相同）。

综上所述，实证研究结果表明，单期收益 h_{1t}、股票价格 p_t 和红利价格比 δ_t' 的波动性一般都很高（即它们的波动性大于相应理论值 h_{1t}'、p_t' 和 δ_t' 的波动性），以至于理性预期假设下的理性定价公式不成立。这一结论对预期收益的各种假设都成立。

尽管短期（如 1 个月、1 年）收益几乎无法预测，但长期（如 3 ~ 5 年）收益是可预测的。然而，似乎可能的情形是，股票的实际价格和由理性定价公式给出的理论价格之间存在一个较大的持久差异。因此，在几种均衡收益模型中，Var 给出的证据会频繁地拒绝有效市场 - 理性定价公式（EMH - RVF）。

12.3 持续性与波动率

这一节里，我们说明如何用 Var 来考察单期收益可预测性与持续性之间的关系及其对股票价格波动率的含义。我们已经注意到，月度收益的可预测性很弱，单方程回归的 R^2 只有 0.02 左右，单变量模型中的持续性是用自回归系数与 1 的接近程度来进行度量的。这一节将说明，如果预期单期收益基本上不可预测但具有持续性，那么收益方面的消息仍然会对股票价格产

生很大的影响。另外，使用 Var 系统可以同时对红利方面的消息、未来收益（贴现率）方面的消息及两者的交互作用对股票价格变化的贡献进行考察。

持续性与消息

Campbell（1991）考虑了股票价格对预期未来贴现率（必要收益）变化的影响和对预期未来红利变化的影响。单期预期收益的预测误差可以表示为（见附录）

$$h_{t+1} - E_t h_{t+1} = (E_{t+1} - E_t)\sum_{j=0}^{\infty}\rho^j \Delta d_{t+j+1} - (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j h_{t+j+1} \tag{42}$$

紧凑形式的表示是

$$v^h_{t+1} = \eta^d_{t+1} - \eta^h_{t+1} \tag{43}$$

$t+1$ 期预料不到的收益 = 未来红利增长方面的消息 - 未来预期收益方面的消息

注意，（43）的左边表示预料不到的资本利得 $p_{t+1} - E_t p_{t+1}$，右边的项 η^d_{t+1} 和 η^h_{t+1} 表示"预期修正"的 DPV。（43）简单地陈述了如下结论：事后收益 h_{t+1} 大于 $E_t h_{t+1}$ 的可能原因是对未来红利增长 Δd_{t+j} 的预期出现向上修正或贴现率 h_{t+j} 的预期出现了向下修正。如果红利增长和贴现率的预期修正具有持续性，那么［t，$t+1$］期间与两者有关（或其中一个有关）的消息对预料不到收益 $h_{t+1} - E_t h_{t+1}$ 及方差有着实质性影响。（42）的右边是两个随机变量 Δd_{t+j+1} 和 h_{t+j+1} 的加权和，预料不到收益的方差可以写成

$$\mathrm{Var}\,(v^h_{t+1}) = \mathrm{Var}\,(\eta^d_{t+1}) + \mathrm{Var}\,(\eta^h_{t+1}) - 2\mathrm{Cov}\,(\eta^d_{t+1},\ \eta^d_{t+1}) \tag{44}$$

正如我们在下面将看到的那样，一旦对 h_{t+1}（及影响 h_{t+1} 的其他变量）的随机过程形式进行了假定（也就是说，满足 Var 系统），我们就可以将预料不到收益方面的消息分解为（44）右边的几个"消息"。Campbell 给出了一种预期收益持续性的度量指标

$$P_h = \sigma\,(\eta^h_{t+1})\ /\sigma\,(u_{t+1}) \tag{45}$$

u_{t+1} 是 $t+1$ 时刻预期收益的创新（innovation），$u_{t+1} = (E_{t+1} - E_t)\ h_{t+2}$，因此 u_{t+1} 是对预期在一个时期上的修正。因此，P_h 可以定义为

$$P_h = \frac{\text{未来收益 DPV 方面消息的标准差}}{\text{单期预期收益方面消息的标准差}}$$

P_h 也许可以进行这样的解释：由（45）可以看出单期预期收益 u_{t+1}1% 的创新会导致未来贴现率 η^h_{t+j} 产生 P_h% 的变化，由（43）可知，这又会导致 P_h% 预料不到的资本损失。

单变量情形

为了说明持续性在解释股票价格变动方面的重要性，考虑一种简单情形是非常有用的。假设预期收益服从 AR（1）过程

$$E_{t+1} h_{t+2} = \beta E_t h_{t+1} + u_{t+1} \tag{46}$$

持续性取决于 β 与 1 的接近程度。可以证明（见附录）

$$P_h = \rho/\ (1 - \rho\beta)\ \approx 1/(1 - \beta) \qquad (\rho \approx 1) \tag{47}$$

$$\mathrm{Var}(\eta^h_{t+1})/\mathrm{Var}(v^h_{t+1}) \approx [(1+\beta)/(1-\beta)]R^2/(1-R^2) \tag{48}$$

R^2 等于股票收益 h_{t+1} 的方差中可预测部分所占的比例。如果 β 接近 1，那么 P_h 就会很大，这意味着持续性很强。

现在，我们可以用（48）来说明，即使单期股票收益 h_{t+1} 的可预测性很低（即 R^2 很低），只要预期收益具有持续性，那么未来收益方面的消息就会对股票价格产生很大的影响。取 $\beta = 0.9$，将单期收益预测方程中的 R^2 取为 0.025，由（48）就得到

$$\mathrm{Var}\,(\eta^h_{t+1}) = 0.49\mathrm{Var}\,(v^h_{t+1}) \tag{49}$$

由（42）和（43）有

$$\mathrm{Var}\,(p_{t+1} - E_t p_{t+1}) \equiv \mathrm{Var}\,(v^h_{t+1})$$

在这种情形中，未来收益方面消息的方差 Var（η^h_{t+1}）可以解释单期预料不到收益方差

Var（v_{t+1}^h）的 49%，从而可以解释股票价格变动的 49%。因此，股票收益的可预测性很弱，如果收益的持续性很强，那么与收益方面的消息仍然可以解释股票价格变动的很大一个部分。

Campbell 使用多变量 Var 系统来对上述单变量模型进行推广，使用的变量有：用市值加权的 NYSE 指数的月度实际收益、红利价格比 δ_t 和国库券的相对利率 rr_t（后者定义为短期国库券利率与其往后一年的移动平均值之差：移动平均部分可以"消除" I（1）利率序列中的趋势）。包含三个变量 $z_t =（h_t, \delta_t, rr_t）$ 的 Var 的伴随形式是：

$$z_{t+1} = Az_t + w_{t+1} \tag{50}$$

这里的 w_{t+1} 是预测误差（$z_{t+1} - E_t z_{t+1}$）。现在，我们可以用（50）和（42）通过"Var 圈"（Var hoops）将预料不到的股票收益 v_{t+1}^h 的方差分解为预期红利方面的消息和预期贴现率（收益）方面的消息。使用 Var 就可以得到（见附录）

$$\begin{aligned} \eta_{t+1}^h &= (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j h_{t+1+j} \\ &= \mathrm{e}1' \sum_{j=1}^{\infty} \rho^j A^j \mathrm{w}_{t+1} \\ &= \rho\ (\mathrm{I} - \rho \mathrm{A})^{-1} \mathrm{e}1' A \mathrm{w}_{t+1} \end{aligned} \tag{51}$$

因为 v_{t+1}^h 是 w_{t+1} 的第一个分量（即 $\mathrm{e}1'\mathrm{w}_{t+1}$），可以将（51）进行重写并根据如下恒等式来计算 η_{t+1}^d：

$$\eta_{t+1}^d = v_{t+1}^h + \eta_{t+1}^h = (\mathrm{e}1' + \rho\ (1 - \rho A)^{-1} \mathrm{e}1' A)\ \mathrm{w}_{t+1} \tag{52}$$

给定 Var 中 A 矩阵的估计值和预测误差的方差协方差矩阵 $\Psi = E\ (w_{t+1} w'_{t+1})$ 的估计值，我们就能计算出方差分解中的各个方差和协方差。持续性指标 P_h 为（见附录）

$$P_h = \sigma\ (\lambda' w_{t+1})\ / \sigma\ (e1' A w_{t+1}) = [\ (\lambda' \Psi \lambda)\ /\ (e1' A \Psi' A' e1)\ ]^{1/2} \tag{53}$$

其中，$\lambda' \rho\ (I - \rho A)^{-1} e1' A$，$\sigma$（.）是括号中项的标准差。

结果

表 3 取自 Campbell（1991）对 1952（1）～1988（12）月度数据的研究结果。在收益的 Var 方程中，h_{t+1} 受滞后变量 h_t 的影响（即受红利价格比和国库券的相对利率的影响），尽管如此，相对于红利收益（$R^2 = 0.96$）和国库的相对券利率（$R^2 = 0.55$）而言，0.065 的 R^2 还是很低。持续性指标的计算结果是 $P_h = 5.7$（$s.e. = 1.5$），这意味着在其他条件不变的情况下，预期收益创新每增加 1% 就会导致约 6% 的资本损失。（然而，1927（1）～1951（12）期间的持续性更低，$P_h = 3.2$，$s.e. = 2.4$。）未来收益方面的消息能解释预料不到收益方差的 75%，红利方面的消息能解释 13%（见表 3），而剩下的 10% 应归因于负的协方差项。

表 3　实际股票收益的方差分解：美国数据，1952（1）～1988（12）

R_h^2	Var（η^d）/Var（v^h）	Var（η^h）/Var（v^h）	-2Cov（η^d, η^h）	Corr（η^h, η^k）
0.065	0.127	0.772	0.101	−0.161
[0.000]	(0.016)	(0.164)	(0.153)	(−0.256)

注：R^2 是月度实际股票收益方差中能用 Var 来进行预测的部分，[.] 是 Var 预测变量的联合置信水平，(.) 等于标准差。Var 的滞后长度为 1。

资料来源：Campbell（1991），表 2。

Campbell（1991）注意到，这些 Var 结果还需要进一步分析，他考虑了 Var 滞后阶数对方差分解的敏感性、Var 中哪些变量可以删去、红利价格方程中的单位根问题，这些都可能对渐近检验统计量的小样本性质产生影响。他发现了如下现象：

（ⅰ）不管使用的是月度收益数据还是季度收益数据，对 Var 不同的滞后阶数而言，上述结果基本上没有什么变化。

（ⅱ）将红利价格比从 Var 中去掉后，未来收益方面消息的方差的重要性大大降低（未来

红利显得更为重要）。

（iii）对 h_{t+1} 进行蒙特卡罗试验（h_{t+1} 与其他变量独立，红利价格比服从单位根过程），试验结果表明，在方差分解统计量上具有“灾难性影响”。在人工生成的序列中 h_{t+1} 不可预测，预料不到股票收益只对未来红利方面的消息有所反应。因此，Var（η_{t+1}^{d}）/Var（υ_{t+1}^{h}）应等于 1，收益方程的 R^2 应等于 0。但在 1927—1980 年整个样本期的蒙特卡罗模拟中，这些结论明显不成立（尽管对战后时期而言这些假设并没有遭到拒绝）。

Cuthbertson 和 Galindo（1999）使用英国 1918—1993 的年度市值加权 BZW 股票指数数据重复了 Campbell 的分析，他们将更多的变量增加到 Var 模型中，关键变量有红利价格比和对波动率度量指标。他们所发现的一些证据表明，波动持续性有助于解释预期收益的持续性。然而，就对当前收益变动的贡献而言，未来收益方面的消息是与红利方面消息的 4 倍（协方差项也是统计上显著的），这些结果与 Campbell（1991）对美国数据的研究结果基本一致。

Campbell（1991）使用美国股票市场总指数的研究结果和 Cuthbertson、Hayes 和 Nitzsche（1997，1999）的研究结果都发现了多变量 Var 系统中的股票收益（特别是 20 世纪 50 年代后的数据）具有（弱）可预测性和合理的持续性。当然，这些分析无法提供一个经济模型来解释为什么预期收益 $E_t h_{t+1}$ 取决于红利价格比和国库券相对利率等变量，只是给出一堆有待进一步解释的统计相关性。Campbell（1991）及 Cuthbertson 和 Galindo（1999）也发现，尽管两者的相对重要性难以精确确定，股票价格的变动主要应归因于未来收益方面的消息而不是未来红利增长率方面的消息。这是因为，方差分解的结果依赖于选择的特定信息集合（及红利是否是单位根）。

12.4 小结

- Var 方法和 RVF 的线性化允许我们在一个很一般的理论框架下考察单期收益、多期收益和股票价格波动率之间的关系，这个框架也会涉及数据的平稳性问题。
- 在单期收益决定因素的各种假设下，现有证据表明股票价格不满足 RVF，RE 下的信息效率假设也不成立。有效市场 - 理性定价公式（EMH - RVF）遭到拒绝所得出的结论似乎比方差边界文献所得到结果的稳健性更强。即使我们考虑贴现率和红利预测值的时变性质（使用线性 Var 模型），股票价格仍然存在过度波动性。
- 尽管月度收益很难预测，Var 方法还是说明长期收益可预测（这与收益单方程回归的结果一致）。
- 一些证据表明，板块收益取决于板块自身的方差，而不是板块收益与市场收益之间的协方差（即 CAPM 理论）。这可以解释为投资者的“狭窄框架”或股票分析师的市场分割。
- 单期收益存在一定的持续性，尽管单期收益几乎无法预测，当前收益方面的消息还是对未来收益和股票价格有着很大的影响。股票价格很有可能受预期未来收益变动的影响，而不受红利预测值变动的影响（但两者所起的作用相同）。

附录：收益、方差分解与持续性

在本章的附录中，我们为读者介绍三个方面的内容。首先说明如何得出 Campbell - Shiller 关于股票收益和红利价格比的线性方程；接下来看看如何用这些方程来进行 Campbell 方差分解，讨论这些方程对股票价格波动持续性的重要意义；最后介绍在实证研究中如何使用 Var 方法来估计持续性。

1. 收益的线性化

单期实际持有期收益为

$$H_{t+1}=\frac{P_{t+1}-P_t+D_{t+1}}{P_t} \tag{A1}$$

P_t 表示 t 期末的实际股票价格，D_{t+1}是第 $t+1$ 期的实际红利支付。（股票价格和红利都用某个价格水平来进行平减，如消费物价指数。）我们将（1 加上）实际持有期收益的自然对数记为 h_{t+1}，

$$h_{t+1}=\ln(1+H_{t+1})=\ln(P_{t+1}+D_{t+1})-\ln(P_t) \tag{A2}$$

如果用小写字母表示对数值，那么（A2）就变成

$$h_{t+1}=\ln[\exp(p_{t+1})+\exp(d_{t+1})]-p_t \tag{A3}$$

（A3）中的第一项是p_{t+1}和d_{t+1}的非线性函数，我们在 P 和 D 的几何平均值附近使用一阶 Taylor 级数展开来对其进行线性化：

$$\ln[\exp(p_{t+1})+\exp(d_{t+1})]=k+\rho p_{t+1}+(1-\rho)d_{t+1} \tag{A4}$$

其中，

$$\rho=\bar{P}/(\bar{P}+\bar{D}) \tag{A5}$$

因此，ρ 是比 1 略小的数，k 是常数。使用（A3）和（A4）就可以得到

$$h_{t+1}=k+\rho p_{t+1}+(1-\rho)d_{t+1}-p_t \tag{A6}$$

在（A6）中加减 d_t，定义 $\delta_t=d_t-p_t$ 为对数红利价格比（log dividend - price ratio），我们有

$$h_{t+1}=k+\delta_t-\rho\delta_{t+1}+\Delta d_{t+1} \tag{A7}$$

方程（A7）可以解释为 δ 的线性向前差分方程：

$$\delta_t=-k+\rho\delta_{t+1}+h_{t+1}-\Delta d_{t+1} \tag{A8}$$

使用向前递归代入法来对（A8）进行求解，假设横截性条件成立：

$$\delta_t=\sum_{j=0}^{\infty}\rho^j[h_{t+j+1}-\Delta d_{t+j+1}]-k/(1-\rho) \tag{A9}$$

方程（A9）表明，对数红利价格比等于所有未来收益的贴现值之和，减去所有未来红利的贴现值之和，再减去一个常数。如果当前股票价格过低，从而红利价格比较高，那么要么未来期望收益 h_{t+1}较高，要么红利增长率 Δd_{t+j}较低，或者两者同时成立。方程（A9）是个恒等式，对实际数据而言精确成立。然而，在 t 期末可获得信息的条件下对（A9）的两边求数学期望（我们可以将其视为一种事前关系）：

$$\delta_t=\sum_{j=0}^{\infty}\rho^j[E_t(h_{t+j+1})-E_t(\Delta d_{t+j+1})]-k/(1-\rho) \tag{A10}$$

注意，δ_t 在 t 期末就是已知的，因此它的数学期望就等于它本身。

2. 方差分解

注意，我们可以将 $t+1$ 期的（A10）写成

$$\delta_{t+1}=\sum_{j=0}^{\infty}\rho^j[E_{t+1}(h_{t+j+2})-E_{t+1}(\Delta d_{t+j+2})]-k/(1-\rho) \tag{A11}$$

根据（A8），我们有

$$h_{t+1}-\Delta d_{t+1}-k=\delta_t-\rho\delta_{t+1} \tag{A12}$$

将（A10）和（A11）代入并进行重新排列后就得到

$$\begin{aligned}h_{t+1}-\Delta d_{t+1}-k=&E_t h_{t+1}+\sum_{j=1}^{\infty}\rho^j E_t(h_{t+j+1})-\sum_{j=0}^{\infty}\rho^j E_t(\Delta d_{t+j+1})-k/(1-\rho)\\&-\sum_{i=0}^{\infty}\rho^{i+1}E_{t+1}(H_{t+i+2})+\sum_{i=0}^{\infty}\rho^{i+1}E_{t+1}(\Delta d_{t+i+2})+k\rho/(1-\rho)\end{aligned} \tag{A13}$$

借助 k 和 ρ 将（A13）中的常数项消掉。将 $j=i+1$ 代入（A13）右边最后两个和式中并进行重

新排列就得到

$$h_{t+1}-E_t h_{t+1}=\sum_{j=1}^{\infty}\rho^j E_t\ (h_{t+j+1})\ -\sum_{j=0}^{\infty}E_t\ (\Delta d_{t+j+1})$$
$$-\sum_{j=1}^{\infty}\rho^j E_{t+1}\ (h_{t+j+1})\ +\sum_{j=0}^{\infty}\rho^j E_{t+1}\ (\Delta d_{t+j+1})\ +\Delta d_{t+1} \qquad (A14)$$

对（A14）进行重新排列，我们就得到预料不到的收益或异常收益的关键表示

$$h_{t+1}-E_t h_{t+1}=\ (E_{t+1}-E_t)\ \sum_{j=0}^{\infty}\rho^j\Delta d_{t+1+j}-\ (E_{t+1}-E_t)\ \sum_{j=1}^{\infty}\rho^j h_{t+1+j} \qquad (A15)$$

Campbell（1991）使用方程（A15）来分析预期未来收益 $h_{t+1}-E_t h_{t+1}$ 对预料不到的当前收益的持续性影响，（A15）中的每一项都可以写成

$$v^h_{t+1}=\eta^d_{t+1}-\eta^h_{t+1} \qquad (A16)$$

（$t+1$ 期预料不到的收益）＝（未来红利增长方面的信息）－（未来预期收益方面的信息）

根据（A16），我们有

$$\mathrm{Var}\ (v^h_{t+1})\ =\mathrm{Var}\ (\eta^d_{t+1})\ +\mathrm{Var}\ (\eta^h_{t+1})\ -2\mathrm{Cov}\ (\eta^d_{t+1},\ \eta^d_{t+1}) \qquad (A17)$$

在（A17）中，预料不到的股票收益的方差可以分解为三个部分：现金流（红利）方面信息的方差、未来收益方面信息的方差和协方差项。在这个方差分解中，计算三个部分对股票收益变动的相对重要性是可能的。由（A6）可知，当 $\rho\approx 1$ 且没有预料不到的红利收益时，

$$h_{t+1}-E_t h_{t+1}=p_{t+1}-E_t p_{t+1}$$

p_t 是（对数）股票价格，因此（A15）的左边是预料不到的或"意料之外"的股票价格。

Campbell 也给出了对预期收益持续性进行度量的指标，指标定义为未来收益预期现值的变化（即 η^h_{t+1} 的标准误差）与下一期预期收益变化之比。如果我们定义 u_{t+1} 为 $t+1$ 时刻后面一期预期收益的创新（innovation），那么

$$u_{t+1}=\ (E_{t+1}-E_t)\ h_{t+2} \qquad (A18)$$

预期收益持续性的度量指标 P_h 定义为

$$P_h=\sigma\ (\eta^h_{t+1})\ /\sigma\ (u_{t+1}) \qquad (A19)$$

A. 预期收益服从 AR（1）过程

为了进行（A17）的方差分解和计算（A19）的持续性指标，我们需要对股票预期收益进行建模。为了说明问题，我们遵循 Campbell 的方法，首先假设预期股票收益服从单变量 AR（1）模型，然后在 Var 表示中用 h_{t+1} 来重复计算过程。预期收益的 AR（1）模型为

$$E_{t+1}h_{t+2}=\beta E_t h_{t+1}+u_{t+1} \qquad (A20)$$

使用（A20）就得到 h_{t+2} 在时刻 t 的预期价值

$$E_t h_{t+2}=\beta E_t h_{t+1} \qquad (A21)$$

我们在这里用到了 $E_t E_{t+1}$ 和 $E_t u_{t+1}=0$。方程（A20）减（A21）就得到

$$(E_{t+1}-E_t)\ h_{t+2}=u_{t+1} \qquad (A22)$$

将（A21）的时间向后延展一期，并在时刻 t 求数学期望，我们有

$$E_t h_{t+3}=\beta E_t h_{t+2}=\beta^2 E_t h_{t+1} \qquad (A23)$$

类似地，

$$E_{t+1}h_{t+3}=\beta E_{t+1}h_{t+2}=\beta\ (\beta E_t h_{t+1}+u_{t+1}) \qquad (A24)$$

（A24）减（A23）就得到

$$(E_{t+1}-E_t)\ h_{t+3}=\beta u_{t+1}$$

对更为一般的情形而言，我们有

$$(E_{t+1}-E_t)\ h_{t+j+1}=\beta^{i-1}u_{t+1} \qquad (A25)$$

在（A15）和（A16）中使用未来收益方面信息的定义和（A25）就得到

$$\eta^h_{t+1}=\sum_{j=1}^{\infty}\rho^j\beta^{j-1}u_{t+1}=\rho u_{t+1}/\ (1-\rho\beta) \qquad (A26)$$

因此，预料不到收益的贴现值的方差是单期预料不到收益的方差的函数：

$$\mathrm{Var}\ (\eta_{t+1}^{h})\ =\ [\rho/\ (1-\rho\beta)]^{2}\mathrm{Var}\ (u_{t+1}) \tag{A27}$$

根据（A27），（A19）中持续性度量指标可以表示为

$$P_{h}=\rho/\ (1-\rho\beta)\ \approx 1/\ (1-\beta)$$

因此，如果β接近1，我们可以将其解释为：AR（1）模型中存在很强的持续性，P_h的值也很大。因为$p_{t+1}-E_{t}p_{t+1}=-\eta_{t+1}^{h}$（$\eta^{d}=0$时），$\eta^{h}=[\rho/\ (1-\rho\beta)]\ u_{t+1}$，所以对AR（1）情形而言有$p_{t+1}-E_{t}p_{t+1}=P_{h}u_{t+1}$。于是，$u_{t+1}$每增加1%就会导致$\eta^{h}$增加$P_{h}\%$，也会导致预料不到的资本损失$P_{h}\%$。

对AR（1）情形而言，我们想说明的是，即使我们只能解释单期收益h_{t+1}变化的很小一个部分（收益很难预测），但是如果收益具有持续性，那么收益方面的信息在解释股票价格波动时显得非常重要。简而言之，预期收益的持续性越强，在解释预料不到的收益$h_{t+1}-E_{t}h_{t+1}$（或预料不到资本利得或损失$p_{t+1}-E_{t}p_{t+1}$）时未来收益方面信息的方差$\mathrm{Var}\ (\eta_{t+l}^{h})$就越重要。定义$R^2$为股票收益方差可预测部分所占的比例

$$R^{2}=\mathrm{Var}\ (E_{t}h_{t+1})\ /\mathrm{Var}\ (h_{t+1}) \tag{A28}$$

$$1-R^{2}=\mathrm{Var}\ (v_{t+l}^{h})\ /\mathrm{Var}\ (h_{t+1}) \tag{A29}$$

$$R^{2}/\ (1-R^{2})\ =\mathrm{Var}\ (E_{t}h_{t+1})\ /\mathrm{Var}\ (h_{t+1}^{h}) \tag{A30}$$

根据（A20），$E_{t}h_{t+1}$的方差为

$$\mathrm{Var}\ (E_{t}h_{t+1})\ =\mathrm{Var}\ (u_{t+1})\ /\ (1-\beta^{2}) \tag{A31}$$

将（A31）代入（A30）并对$\mathrm{Var}\ (u_{t+1})$求解就得到

$$\mathrm{Var}\ (u_{t+1})\ =\ (1-\beta^{2})\ \mathrm{Var}\ (v_{t+1}^{h})\ R^{2}/\ (1-R^{2}) \tag{A32}$$

应用（A32），（A27）可以写成

$$\begin{aligned}\mathrm{Var}\ (\eta_{t+1}^{h})\ /\mathrm{Var}\ (v_{t+1}^{h})\ &=\ (1-\beta^{2})\ [\rho/\ (1-\rho\beta)]^{2}/R^{2}\ (1-R^{2})\\&\approx\ [\ (1+\beta)\ /\ (1-\beta)]\ R^{2}/\ (1-R^{2})\end{aligned} \tag{A33}$$

（A33）的左边是方差分解中我们感兴趣的部分，它反映的是预期未来收益贴现值的方差与预料不到收益的方差的相对重要性（见方程（A17））。

就月度收益而言，$R^{2}\approx 0.025$的预测方程具有一定的代表性。当$\beta=0.5$，0.7或0.9时，方差比分别等于VR＝0.08，0.18或0.49。因此，对高持续性和低预测性的收益而言，未来收益方面的信息仍然对预料不到收益的方差v_{t+l}^{h}有很大（比例）的影响。

B. 多变量情形：Var模型

上述单变量情形忽略了预期收益信息与预期红利信息之间的交互作用，即（A17）中的协方差项。我们至少需要一个方程来解释红利增长，然后对红利增长预测误差（即信息）与收益预测误差之间的协方差进行考察。

这一节假设（$m\times 1$）向量z_{t+1}以h_{t+1}作为第一个分量，z_{t+1}中的其他分量在$t+1$期末都是已知的，我们可以构建如下Var模型，为它的首要成分一方程可以解释相互作用：

$$z_{t+1}=Az_{t}+w_{t+1}\quad 且\quad E\ (w_{t+1}w_{t+1}^{'})\ =\Psi \tag{A34}$$

A是伴随矩阵，w_{t+1}中的第一个元素是v_{t+1}^{h}。首先要注意的是

$$E_{t+1}z_{t+j+1}=A^{j+1}z_{t}+A^{j}w_{t+1} \tag{A35}$$

$$E_{t}z_{t+j+1}=A^{j+1}z_{t} \tag{A36}$$

（A35）减去（A36）后就得到

$$(E_{t+1}-E_{t})\ z_{t+j+1}=A^{j}w_{t+1} \tag{A37}$$

因为z_t的第一个分量是h_t，所以如果对（A37）的两边都左乘$E1'$（e1′是（$1\times m$）行向量，第一个分量为1，其余分量为0），我们就得到

$$(E_{t+1}-E_{t})\ h_{t+j+1}=\mathrm{e}1'A^{j}w_{t+1} \tag{A38}$$

因此

$$
\begin{aligned}
\eta_{t+1}^{h} &= (E_{t+1}-E_{t})\sum_{j=1}^{\infty}\rho^{j}h_{t+1+j} \\
&= \mathrm{e1}'\sum_{j=1}^{\infty}\rho^{j}A^{j}w_{t+1} \qquad (\mathrm{A39}) \\
&= \mathrm{e1}'\rho A\,(I-\rho A)^{-1}w_{t+1} \\
&= \boldsymbol{\lambda}'w_{t+1}
\end{aligned}
$$

$\lambda'=\mathrm{e1}'\rho A\,(I-\rho A)^{-1}$是 Var 参数的非线性函数。因为$w_{t+1}$的第一个分量是$v_{t+1}^{h}$，所以使用（A16）和（A39）就得到

$$\eta_{t+1}^{d}=\mathrm{e1}'\,[I+\rho A\,(I+\rho A)^{-1}]\,w_{t+1}=\gamma w_{t+1} \qquad (\mathrm{A40})$$

从（A39）和（A40）可以看出，预料不到的未来收益和预料不到的未来红利都可以写成 Var 误差项的线性组合，每一个误差项都要乘以 Var 参数的一个非线性函数。在（A38）中令$j=1$并使用（A18）就得到

$$(E_{t+1}-E_{t})\,h_{t+2}=u_{t+1}=e1'Aw_{t+1} \qquad (\mathrm{A41})$$

给定 Var 模型，方程（A39）～（A40）可以用来进行方差分解和计算持续性指标P_h。特别地，我们有

$$\mathrm{Var}\,(\eta_{t+1}^{h})=\lambda'\Psi\lambda \qquad (\mathrm{A42})$$

$$\mathrm{Var}\,(\eta_{t+1}^{d})=\gamma'\Psi\gamma \qquad (\mathrm{A43})$$

$$\mathrm{Var}\,(v_{t+1}^{h})=\mathrm{e1}'\Psi\mathrm{e1} \qquad (\mathrm{A44})$$

$$\mathrm{Cov}\,(\eta_{t+1}^{d},\ h_{t+1}^{d})=\lambda'\Psi\gamma \qquad (\mathrm{A45})$$

$$\mathrm{Var}\,(u_{t+1})=\mathrm{e1}'A\Psi A'\mathrm{e1} \qquad (\mathrm{A46})$$

$$P_h=[\,(\lambda'\Psi\lambda)\,/\,(\mathrm{e1}'A\Psi A'\mathrm{e1})\,]^{1/2} \qquad (\mathrm{A47})$$

一旦 Var 的参数“A”和协方差矩阵Ψ被估计出，所需要的方差和协方差也就很容易计算。我们可以用 OLS 来对 Var 的每个方程进行分开估计，但 Campbell 建议使用广义矩（GMM）估计法（Hansen（1982））来对误差项中存在的异方差进行修正。尽管模型参数的 GMM 方差－协方差矩阵可以对异方差进行“修正”，参数的 GMM 点估计还是与 OLS 估计相同（White（1984））。

（A42）～（A47）中方差统计量的标准误差可以按照如下程序来进行计算。用θ表示模型中所有参数构成的向量（包括A和Ψ中的非冗余分量），将这些参数对异方差进行调整后估计值的方差－协方差矩阵记为V。例如，假设我们感兴趣的是计算P_h的标准差。因为P_h是θ的非线性函数，所以它的方差可以用下式来进行计算：

$$\mathrm{Var}\,(P_h)=\left(\frac{\partial P_h}{\partial\theta}\right)'V\left(\frac{\partial P_h}{\partial\theta}\right) \qquad (\mathrm{A48})$$

P_h对参数θ的导数可以用数值方法来进行计算，P_h的标准差就是 Var（P_h）的平方根。

第13章

SDF 模型与 CCAPM 模型

目的

- 叙述一类著名的模型：SDF 模型（随机折现因子模型）。
- 证明以消费为基础的资本资产定价模型（CCAPM 模型）是 SDF 模型，模型中的随机折现因子取决于当前和未来各期消费的边际效用。
- 根据均值方差组合理论来阐述 CCAPM 模型与“标准”Sharpe – Lintner CAPM 模型之间的关系。
- 说明 CCAPM 模型中利率、风险资产收益和资产价格之间是怎样相互联系的。

在下面的部分里，我们讨论以消费为基础的资本资产定价模型（CCAPM 模型），它是比前面章节中所讨论的“标准”均值方差 CAPM 模型更为一般的资产定价框架。在这个模型中，投资者不像标准 CAPM 模型那样在单期收益的均值和标准差基础上来决定自己的行为，这个模型属于跨期模型，模型中的投资者最大化当前和未来消费的预期效用。

在 CCAPM 模型中，金融资产允许消费者匀滑不同时期的消费：售卖资产来为“萧条”时期的消费进行融资，在“繁荣”时期进行储蓄。那些收益与消费之间条件协方差为负且值很大的资产，即使预期收益很低人们也愿意持有。这是因为，在最需要这些资产的时候，即消费很低的时候，这些资产能“转化为现金”，因此，额外消费产生了更高的边际效用。CCAPM 模型将资产的系统风险与经济状态（即消费）联系起来。

SDF 模型是一类模型，模型中的资产收益 $R_{i,t+1}$ 或价格 P_{it} 可以表示为 $E_t\{R_{i,t+1}M_{t+1}\} = 1$ 或 $P_{it} = E_t\{M_{t+1}X_{i,t+1}\}$，$M_{t+1}$ 在这里表示随机折现因子（或定价核），$X_{i,t+1}$ 是下一期的资产偿付（pay off）。M_{t+1} 的不同构造对应的是不同的 SDF 模型，在这一章节里，我们主要考虑 CCAPM 模型。SDF 模型的核心内容是，M_{t+1} 对所有资产而言都是相同的。

CCAPM 模型是 SDF 模型。在 CCAPM 模型中，在幂效用函数的假设下折现率由消费增长来决定。事实上，消费、无风险利率、风险资产的预期均衡收益和资产价格在 SDF 模型中是同时确定的。在 SDF 模型中 M_{t+1} 可能具有时变性，从而时变性风险溢价存在。

13.1 以消费为基础的资本资产定价模型

单期标准 CAPM 模型假设投资者的目标函数完全由（单期）投资组合的标准差和预期收益来确定。所有投资者选择风险资产的份额来最大化夏普比。因为所有投资者都有相同的预期，所有资产都被投资者持有，均衡收益产生。

CCAPM 模型给出了确定均衡收益的另外一种观点。在这个模型中，投资者最大化预期效

用，而预期效用仅仅依赖于当前和未来的消费（参见 Lucas（1978），Mankiw 和 Shapiro（2001），Cochrane（2001））。金融资产在模型中起着重要作用，有助于匀滑不同时期的消费。持有证券的目的在于将购买力从一个时期转移到另一个时期。如果投资者没有任何资产，也不允许他积累资产，那么他的消费就由当前收入来完全决定。如果他持有资产，那么在当前收入很低时，他可以通过变卖这些资产来为消费融资。所以，当预期消费很低时，如果个体资产的预期收益很高，那么这些个体资产就是更为“合意”的资产。因此，资产的系统风险由资产收益与消费之间的协方差（而不是像“标准”CAPM 模型中那样由资产收益与投资组合收益之间的协方差）来确定。

正如我们即将看到的那样，CCAPM 模型有许多等价的表述方式，有些表述方式比其他表述方式更为直观。简而言之，CCAPM 模型的变化比 20 世纪 30 年代由 Busy Berkeey 设计舞蹈动作的非常成功的好莱坞音乐的变化还要多，它在资产定价和投资组合的整体文献中看上去非常规则，模型色彩斑斓且根基牢固。模型本身经历了一些有趣的发展（从 1933 年的掘金者，到具有歌剧幻影的酒店歌舞表演，再到一些纤细的隐喻），目的是为了更好地解释经验事实。在以下几章的末尾部分里，你将自己判断这纯粹是带有娱乐性的理论分支，还是为了更为充分地解释“现实世界”而使其在资产收益和风险溢价分析方面更为有用。

记号

下一期持有资产的美元偿付为 $X_{t+1} \equiv P_{t+1} + D_{t+1}$，$P_{t+1}$是（期末）价格，$D_{t+1}$是 t 与 $t+1$ 之间的红利收益。资产的总收益定义为 $R^*_{t+1} = X_{t+1}/P_t \equiv R_{t+1} + 1$。因此，$R^* = 1.10$ 的总收益（使用通常的表示方式）表示 $R = 0.10$（10%）的“收益”。

一阶条件

我们首先陈述导出投资者一阶条件（first - Order Condition，FOC）的一种直观方式，投资者所关心的问题是如何使（真实）消费的终身效用最大化。（效用函数具有时间分离形式——参见第 1 章。）因为投资者关心的是实际消费，所以资产收益也是指实际收益。最大化预期效用的一阶条件是，投资者当前消费减少所导致的效用损失等于下一期所导致的额外（折现）收益。时刻较低的消费开支允许投资者在预期收益为正的资产上进行投资，为未来消费提供额外的资金资源。更为正式的叙述是，今天少（实际）消费 1 美元将减少 $U'(C_t)$ 的效用，而下一期的预期偿付为 $E_t R^*_{i,t+1}$。下一期消费 1 美元（实际值）的折现效用为 $\theta U'(C_{t+1})$，折现率为 $\theta = 1/(1+t_p)$，其中 t_p 为主观时间偏好率，用年度数据表示的值为 $\theta \approx 0.97$。因此，下一期预期总效用的现值为 $E_t\{R^*_{i,t+1}\theta U'(C_{t+1})\}$。在均衡条件下，我们有

$$U'(C_t) = E_t\{R^*_{i,t+1}\theta U'(C_{t+1})\} \tag{1a}$$

或

$$E_t\{R^*_{i,t+1}M_{t+1}\} = 1 \tag{1b}$$

其中

$$M_{t+1} = \theta U'(C_{t+1})/U'(C_t) \tag{1c}$$

M_{t+1}是当前消费与未来消费（折现）之间的边际替代率，它依赖于投资者在今天消费和明天消费之间的“偏好”（“品味”）。方程（1b）对任意两种资产 i 和 j 都成立，因此，我们能直接导出用超额收益表示的一阶条件

$$E_t\{(R^*_{i,t+1} - R^*_{j,t+1})M_{t+1}\} = 0 \tag{1d}$$

对无风险资产而言，$R^*_{j,t+1} = R^*_{ft}$。注意，无论所使用的收益是名义收益还是实际收益，方程（1d）都不会发生改变，因为实际收益等于名义收益减去通货膨胀率，而通货膨胀率在方程（1d）中没有出现。

两期模型

下面在“扰动”思想基础上给出稍微正式一点的导出方式（Cochrane（2001））。为简单

起见，我们考虑两个时期的情形，效用函数在时间维度上具有可分性，

$$\max U = U(C_t) + E_t[\theta U(C_{t+1})] \tag{2}$$

（这种表述方式似乎有点傲慢，因为我们用 U 在方程的左边用来表示“总效用”（为了避免混淆，我们在这里要引入另外的符号，t 和 $t+1$ 时刻的效用函数分别为 $U(C_t)$ 和 $U(C_{t+1})$），但我们认为这不会引起任何混淆。）约束条件为

$$C_t = \bar{C}_t - P_t N \tag{3a}$$

$$C_{t+1} = \bar{C}_{t+1} + X_{t+1} N \tag{3b}$$

其中：

C_t = 实际非耐用品消费加耐用品上的服务消费；

$\bar{C}_t$ = 资产购买之前的消费水平；

P_t = 资产价格；

N = 资产的购买数量；

$X_{t+1} \equiv P_{t+1} + D_{t+1} = t+1$ 时的偿付；

θ = 主观折现因子 $(0 < \theta < 1)$。

效用函数的时间可分性意味着任何时期的边际效用不依赖于其他时期的消费。例如，$U = (C_t \cdot C_{t+1})^{\delta}$ 就不具有时间可分性，因为 $\partial U / \partial C_t = \delta C_t^{\delta-1} C_{t+1}^{\delta}$，边际效用依赖于 C_t 和 C_{t+1}。（时间）加法可分性意味着 $U = f(C_t) + \theta f(C_{t+1})$，这保证了 $\partial U / \partial C_t$ 只依赖于 C_t 而不依赖于 C_{t+1}。

假设你的消费是最优的且不存在任何金融资产。现在，我们对你实施“扰动”，你现在可以选择一些金融资产。这种“扰动”方式使我们能建立任何资产和任何个体均衡收益的一阶条件。购买 N 单位任何资产将减少当前消费 $P_t N$。然而，如果这 N 单位资产中的每一单位在 $t+1$ 时的偿付为美元 X_{t+1}，那么 $t+1$ 时用于消费的额外资金就由最初的 $\bar{C}_t$ 增加到 $\bar{C}_{t+1} + X_{t+1} N$。

利用约束条件（3a）和（3b）对方程（2）关于进行最大化就得到

$$P_t U'(C_t) = E_t[\theta U'(C_{t+1}) X_{t+1}] \tag{4}$$

对上式进行重新排列就得到对任何资产都成立的核心定价关系：

$$P_t = E_t(M_{t+1} X_{t+1}) \tag{5}$$

对 CCAPM 模型而言随机折现因子为

$$M_{t+1} = \theta U'(C_{t+1}) / U'(C_t) \tag{6}$$

M_{t+1} 也被称为定价核，在 CCAPM 模型中，它就是当前消费与未来消费之间的边际替代率（MRS）。对方程（2）求全微分并令其等于 0 就可以很容易地看出这一点：

$$\mathrm{d}U = U'(C_t)\,\mathrm{d}C_t + \theta U'(C_{t+1})\,\mathrm{d}C_{t+1} = 0 \tag{7}$$

如果终身效用 U 是一常数，那么这意味着

$$\left(\frac{\mathrm{d}C_t}{\mathrm{d}C_{t+1}}\right) = -\frac{\theta U'(C_{t+1})}{U'(C_t)} \tag{8}$$

总收益 R^* 和（净）收益 R 为

$$R^*_{t+1} \equiv 1 + R_{t+1} \equiv \frac{X_{t+1}}{P_t} \tag{9}$$

因此，我们就演绎出如下结论：收益可由这样的方式来进行确定，

$$1 = E_t[M_{t+1} R^*_{t+1}] \tag{10}$$

这是我们在前面曾经获得过的结果。注意，方程（5）和方程（10）对任意个体都是成立的，并不需要效用函数的任何特殊假设（除加法可分性和凹性外），也不需要资产收益分布的任何假设。只有在具体运用模型时我们才使用一些额外的假设。

无风险利率

“无风险”利率由什么确定呢？如果无风险证券可以交易，那么将方程（10）的“R”拿

到数学期望外面就得到

$$R_{ft}^{*}=\frac{1}{E_t\ (M_{t+1})} \tag{11}$$

如果无风险证券无法进行交易，那么方程（11）中的 R_f 就是“影子”无风险利率（或“零 β”利率）。方程（11）不是特别的直观，我们需要将无法观察的SDF用一些可以观察的变量来进行“替换”，一种方式是选择显示效用函数并假设消费增长服从对数正态分布。

幂效用函数和对数正态消费增长

对幂效用函数而言我们有

$$U\ (C_t)\ =\frac{1}{(1-\gamma)}C_t^{1-\gamma} \tag{12}$$

注意，当 $\gamma\to 1$ 时，我们就得到“对数效用”函数 $U\ (C_t)\ =\ln\ (C_t)$。根据方程（6）和（12），

$$M_{t+1}=\theta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \tag{13}$$

一个标准的统计结果是，如果 Z 服从对数正态分布，那么 $\ln Z\sim N\ (\mu_{\ln z},\ \sigma_{\ln z}^2)$，且

$$E\ [kZ]\ =\exp\left[k\mu_{\ln z}+\frac{1}{2}k^2\sigma_{\ln z}^2\right] \tag{14}$$

其中 k 是常数。利用方程（11）、（13）和（14）就可以得到如下形式的无风险利率（参见附录）

$$R_{ft}^{*}=\left[\theta\left\{\exp\left[-\gamma E_t\Delta c_{t+1}+\frac{\gamma^2}{2}\sigma_t^2\ (\Delta c_{t+1})\right]\right\}\right]^{-1} \tag{15}$$

其中 $\Delta c_{t+1}\equiv\ln\ (C_{t+1}/C_t)$。取对数后就得到

$$\ln R_{ft}^{*}\ =\ -\ln\theta+\gamma E_t\Delta c_{t+1}-\frac{\gamma^2}{2}\sigma_t^2\ (\Delta c_{t+1}) \tag{16}$$

注意 $\ln R_{ft}^{*}=\ln\ (1+R_f)\ \approx R_f$。（在连续时间情形中，$\Delta c\equiv \mathrm{d}C/C$，方程（16）中的最后一项不存在。）根据方程（16），我们能演绎出消费增长率和无风险利率之间的如下关系：

（a）当消费增长率很高时，实际利率也很高。为了用今天的储蓄来增加明天的消费，需要高的实际利率来降低今天的消费。

（b）当 θ 低时（给定 $E_t\Delta c_{t+1}$），实际利率却很高。当人们都没有耐心且未来消费的折现比率很大时，需要高的实际利率来鼓励大家储蓄（因此明天就有额外的消费 $E_t\Delta c_{t+1}$）。

（c）假设预期消费增长具有很高的波动率（相对于消费增长的任何水平而言），那么随着消费的减少，你所损失的效用比等量消费增加所导致的效用增加要多。（这是因为效用函数的凹性和效用的边际递减效应。）因此，随着消费波动率的增加，人们愿意进行更多的储蓄，实际利率就会进一步降低。这是谨慎储蓄的一种形式。

（d）如果 $\gamma=0$（即线性效用函数，从而无任何风险规避倾向），那么实际利率是常数且等于主观折现因子 θ。

（e）随着（幂）效用函数凹性 γ 的增加，实际利率对消费增长的反应更为强烈。例如，今天较高的预期消费增长率要求更高的实际利率，因为较高的 γ 意味着人们更不愿意偏离匀滑消费路径。

注意，在方程（16）中消费是内生的，因此我们也能从相反的方向来解释我们的“相关性”。例如，高的实际利率导致了 t 时刻高的储蓄率和 $t+1$ 时刻更高的消费和消费增长率。

风险资产的收益

现在，我们转为讨论随机折现因子模型中风险资产的价格和收益。是什么决定了资产收益的运动呢？相对于风险资产 j 的平均收益而言，是什么决定了风险资产的平均收益呢？这两种

效应的关键决定因素是随机折现因子 M_{t+1} 与资产收益之间协方差的大小。用任意一种风险资产 i 来表示的 CCAPM 模型为（为了记号的方便，我们在这里省略了一些时间下标）：

$$1 = E_t\ (MR_i^*) \tag{17}$$

对任意两个随机变量 x 和 y 而言，有

$$E\ (xy)\ = \mathrm{Cov}\ (x,\ y)\ +\ (Ex)\ (Ey) \tag{18}$$

因此，运用（17）和（18）就得到

$$1 = E_t\ (M)\ E_t\ (R_i^*)\ + \mathrm{Cov}\ (M,\ R_i^*) \tag{19}$$

$$E_t\ (R_i^*)\ = \frac{1}{E_t\ (M)}\ [1 - \mathrm{Cov}_t\ (R_i^*,\ M)] \tag{20}$$

在方程（20）中运用 $R_f^* = 1/E_t\ (M)\ = U'\ (C_t)\ /\theta E_t U'\ (C_{t+1})$ 并注意到

$$\mathrm{Cov}_t\ (M,\ R_i^*)\ = \theta \mathrm{Cov}_t\ [U'\ (C_{t+1}),\ R_i^*]\ /U'\ (C_t)$$

我们就得到确定任意风险资产 i 超额收益的核心方程

$$\begin{aligned}(E_t R_i^* - R_f^*)_{t+1} &= -R_f^* \mathrm{Cov}_t\ (M,\ R_i^*) \\ &= -\frac{\mathrm{Cov}_t\ [U'\ (C_{t+1}),\ R_{i,t+1}^*]}{E_t\ [U'\ (C_{t+1})]}\end{aligned} \tag{21}$$

注意，"C_{t+1}" 在分子和分母中同时出现。两种资产 i 和 j 的相对（预期超额）收益不同的唯一原因是，R_i 或 R_j 与消费（边际效用）之间的协方差不同。资产收益与 $U'\ (C_{t+1})$ 之间的协方差为负（从而与 C_{t+1} 之间的协方差为正）的那些资产将提供高的预期收益，目的是为了使投资者愿意持有资产。这是因为，资产在消费高的时候得到偿付，消费者已经"感觉很好"，因此更高收益给他们带来的额外效用却非常少。反之，资产收益与消费之间存在负协方差（与 $U'\ (C_{t+1})$ 之间存在正协方差）的那些资产提供的是一种保险，即使所承诺的预期收益很低，人们都愿意持有这些资产。

如果我们假设消费增长和资产收益服从联合对数正态分布，那么可以证明（参见附录）超额收益为

$$E_t\ (r_{i,t+1}^* - r_{ft}^*)\ + \frac{1}{2}\sigma_t^2\ (r_{i,t+1}^*)\ = -\mathrm{Cov}_t\ (m_{t+1},\ r_{i,t+1}^*)$$

其中 $r_{i,t+1}^* = \ln R_{i,t+1}^*$。第二项是 Jensen 效应，协方差项为风险溢价。

13.2 CCAPM 模型与"标准"CAPM 模型

SDF 模型的一阶条件（5）或（10）对具有如下偏好的任何个体都成立：效用函数具有时间可分性且只依赖于消费。如果现在假设存在一个代表性投资者（即所有投资者都有相同的效用函数），那么一阶条件可以解释为均衡预期收益的确定方程，从而一阶条件代表的是以消费为基础的资本资产定价模型（CCAPM 模型）。在方程（21）中应用 $R^* = 1 + R$，就得到

$$E_t R_{i,t+1} - R_{f,t} = \frac{-\mathrm{Cov}_t\ \{U'\ (C_{t+1}),\ R_{i,t+1}\}}{E_t\ \{U'\ (C_{t+1})\}} \tag{22}$$

标准 CAPM 模型与 CCAPM 模型之间有什么关系呢？如果由所有风险资产构成的市场投资组合的收益 $R_{m,t+1}$ 与 $t+1$ 时消费的边际效用 $U'\ (C_{t+1})$ 完全相关（这意味着它与 M_{t+1} 之间具有完全的条件相关性），那么标准 CAPM 模型可由 CCAPM 模型导出。市场投资组合有时也被称为财富组合，因为没有劳动收入，所以市场投资组合包含了所有财富。为了记号更为简便，U'_{t+1} 与 R_m 之间的完全线性关系可以表示为：

$$U'_{t+1} = a + bR_m \tag{23}$$

我们在这里用下标 m 来表示均值方差模型中的最优投资组合。由（23）可以得到

$$\mathrm{Cov}_t\ \{U'_{t+1},\ R_i\}\ = b\mathrm{Cov}\ (R_i,\ R_m) \tag{24a}$$

$$\mathrm{Cov}_t\ \{U'_{t+1},\ R_i\} = b\mathrm{Var}\ (R_m) \tag{24b}$$

$$\mathrm{Var}_t\ (U'_{t+1}) = b^2\mathrm{Var}\ (R_m) \tag{24c}$$

因此

$$\rho\ (U'_{t+1},\ R_m) = \frac{\mathrm{Cov}_t\ (U'_{t+1},\ R_m)}{\sqrt{\mathrm{Var}\ (U'_{t+1})}\sqrt{\mathrm{Var}\ (R_m)}} \tag{24d}$$

将（22）用于市场投资组合（即用 R_m 来代的 R_i）并进行重新排列后就得到

$$\frac{1}{E_t\ \{U'\ (C_{t+1})\}} = \frac{-\ (E_tR_m - R_f)}{\mathrm{Cov}_t\ (U'_{t+1},\ R_m)} \tag{25}$$

将（25）中的 $E_t\ \{U'\ (C_{t+1})\}$ 代入（22），然后利用（24a）和（24b），我们就得到标准 CAPM 模型

$$E_t\ (R_i - R_f) = \beta_t E_t\ (R_m - R_f) \tag{26a}$$

其中

$$\beta_t = \frac{\mathrm{Cov}_t\ (R_i,\ R_m)}{\mathrm{Var}_t\ (R_m)} \tag{26b}$$

因此，标准 CAPM 模型可由 CCAPM 模型导出，我们并没有对资产收益的分布进行任何特殊假设。然而，我们不得不做出可能不现实的假设，消费的边际效用与市场投资组合收益之间完全负（条件）相关——因此我们在这里获得了一个有点比鲁斯王式的（Pyrrhic）胜利。我们已经说明了 CCAPM 和标准 CAPM 模型之间的关系，一般而言，两者是不一致的。例如，如果效用函数是二次函数，那么边际效用是消费的线性函数，我们的消费意味着 C_{t+1} 和 $R_{m,t+1}$ 是完全条件相关的，但数据不可能支持这一结果。最后要注意的是，如果我们假设 SDF 为 $M_{t+1} = a_t + b_tR_{m,t+1}$，那么上述所有结果都成立。这种情形的出现是因为 M_{t+1} 中的分母为 $U'\ (C_t)$，它是非随机的（以时刻形成的预期为条件），时变的 $\{a_t,\ b_t\}$ 在 t 时刻就已经知道了。

可以证明线性关系 $M_{t+1} = a_t + b_tR_{m,t+1}$ 可以由如下假设导出（Cochrane（2001））：

1. C_t 和 C_{t+1} 的两期二次效用函数；
2. 单期指数效用函数和多维正态收益；
3. 对数效用函数和正态分布收益；
4. 无限期界上具有时间可分性的二次效用函数和独立同分布收益。

注意，上述假设中的任何一个都不允许在最大化问题中出现劳动收入，情形（1）和（2）可以导出条件 CAPM 模型，因为 a_t 和 b_t 依赖于 t 时刻的一些变量（如 C_t 和 $R_{m,t}$）。

两期问题

考虑两期二次效用函数这一最简单的情形：

$$\max\left\{-\frac{1}{2}\ (C_t - C)^2 - \frac{1}{2}\theta E_t\ (C_{t+1} - C)^2\right\}$$

约束条件为

$$C_{t+1} = W_{t+1} = R_{m,t+1}\ (W_t - C_t)$$

$$R_{m,t+1} = \sum_{i=1}^{n} \alpha_i R_{i,t+1}$$

$$\sum_{i=1}^{n} \alpha_i = 1$$

其中，$R_{m,t+1}$ 表示总收益（为了与第 15 章的内容相一致），C 是常数，预算约束意味着“市场投资组合”的收益就是财富 W_t（通常是外生给定的）收益。投资者选择 C_t、C_{t+1} 和投资组合在 n 种资产上的权重 α_i。SDF 为

$$M_{t+1} = \frac{\theta U'\ (C_{t+1})}{U'\ (C_t)} = \frac{\theta\ (C_{t+1} - C)}{C_t - C}$$

用预算约束方程中的 C_{t+1} 来进行代替后就得到

$$M_{t+1}=\frac{-\theta C}{(C_t-C)}+\theta\frac{(W_t-C_t)}{C_t-C}R_{m,t+1}$$

因此，M_{t+1} 是 $R_{m,t+1}$ 的线性函数（以 W_t 和 C_t 为条件）。

考虑上述第二种情形。我们已经在单期负指数效用函数和第 1 章中多维正态变量假设的基础上导出了资产需求函数，这里的 $C_{t+1}=W_{t+1}$，因此

$$U(W_{t+1})=-e^{-bW_{t+1}}=-e^{-bC_{t+1}}$$

预算约束为 $W_{t+1}=C_{t+1}=A_fR_f+A'R_{t+1}$，$R_{t+1}$ 为风险资产向量，$W_t=A_f+A'e$。持有风险资产的最优数量 $A=(A_1, A_2, \cdots, A_n)$ 为

$$A=b^{-1}\Sigma^{-1}[E(R)-R_fe]$$

其中，Σ 是 n 种风险资产收益 R 的（条件）方差 - 协方差矩阵，b 为常相对风险规避系数，e 是数值全为 1 的向量。注意，A 为持有风险资产的数量，它独立于财富数量。（无风险资产的需求为 $A_f=W_t=\sum_{i=1}^{n}A_i$。）如果所有投资者都“相同”，那么

$$ER-R_fe=b\sum A=b\mathrm{Cov}(R, R_m)$$

其中，ΣA 给出的是每种收益与 $R_m=A'R+A_fR_f$ 之间的协方差，因此它也给出了投资组合中资产收益之间的协方差。因此，在期末财富 W_{t+1}（或 C_{t+1}）的负指数效用函数假设下，CCAPM 模型给出的是标准 CAPM 模型，因为消费再一次由市场（财富）投资组合来代替。

多期问题

上述“第三种”情形假设效用函数为对数效用函数 $U(C)=\ln C$。在第 15 章中我们将证明，在跨期消费 - 投资组合模型中对数效用函数给出的最优消费与财富成比例（假设不存在随机的劳动收入）。因此，财富组合的收益与消费增长和 SDF 成比例，M 是财富组合收益的线性函数。对数效用函数是另外一种特殊情形，预期收益变化的收入效应和替代效应刚好相互抵消，从而 C/W 独立于收益。

最后，我们考虑第四种情形。正如我们即将在后续几章中要见到的那样，价值函数 $V(\cdot)$ 的概念允许你将多期优化问题改写为两期优化问题，因为

$$V(W_t)=\max_{(c_t,\alpha_t)}\{U(C_t)+\theta E_t[\max_{C_{t+1},C_{t+1},\cdots}^{\alpha_{t+1},\alpha_{t+2},\cdots}E_{t+1}\sum_{j=0}^{\infty}\theta^jU(C_{t+j+1})]\}$$

可以写成

$$V(W_t)=\max_{(c_t,\alpha_t)}\{U(C_t)+\theta E_tV(W_{t+1})\}$$

其中，$V(\cdot)$ 为价值函数，“表示”由所有时期中最优消费（C_{t+1}，C_{t+2}，…）导出的效用函数。为了强调这一点，注意价值函数 $V(W_{t+1})=V[R_{m,t+1}(W_t-\tilde{C}_t)]$，这里的 $\tilde{C}_t$ 是为最优消费水平（求解一阶条件就可以得到最优消费水平）。因此，如果你将你的财富进行最优投资，那么价值函数 $V(W_t)$ 是已实现的期望效用水平，价值函数是财富的函数，因为财富越多你能获得的潜在幸福感就越强。$V(\cdot)$ 也可能是其他变量的函数，我们将在下面及后面的章节中讨论这种可能性。

这些方程在数学上“讲述了一个相同的故事”。第一个方程说的是，你要最大化一生的折现效用就不得不考虑（例如）今天一份额外的比萨饼所带来的边际效用增加和 10 年时间不多喝一瓶酒所带来的效用损失。第二个方程的解释更加吸引人。比如说，你要对今天的一个假期所带来的边际效用增加进行评价，因为今天的假期会导致明天财富减少从而导致预期效用损失 $V'(W_{t+1})$（不是针对 $U'(C_{t+1})$（$j=1, 2, \cdots$）的效应，$U'(C_{t+1})$ 是降低消费所导致的未来各期的边际效用损失，例如后面四年或更长的时间内不去看电影）。可以证明，如果收益独立同分布，那么在约束条件 $W_{t+1}=R_{m,t+1}(W_t-C_t)$ 下最大化关于消费的跨期二次效应函数所得到的一阶条件与两期情形类似：

$$P_tU'(C_t)=\theta E_t[V'(W_{t+1})X_{t+1}]\quad \text{或}\quad E_t(M_{t+1}R_{m,t+1})=1$$

其中

$$M_{t+1}=\theta\frac{V'(W_{t+1})}{U'(C_t)}=\theta\frac{V'(W_{t+1})}{U'(W_t)}$$

$V(\cdot)$ 为价值函数，$U'(C_t)=V'(W_t)$ 是包络条件。注意，$U'(C_t)=V'(W_t)$ 是因为，在最优点，今天消费所带来的额外边际效用必须等于财富的任何变化所带来的边际效用改变。我们将在第 15 章中更为宽泛地讨论价值函数，但现在我们应注意，如果关于消费的跨期效用函数是二次函数（依赖于 $(C_{t+j}-C)^2$ 项），那么在预算约束条件（无劳动收入）和独立同分布收益假设下，价值函数是关于财富的二次函数。然而，就一般情况而言，任何（可接受的）形式的（如幂函数）效用函数都不能使价值函数 $V(W_t)$ 具有相同的函数形式。因此，如果这一“神秘”的价值函数是关于财富的二次函数，那么

$$V(W_{t+1})=-\frac{\phi}{2}(W_{t+1}-W)^2$$

$$M_{t+1}=\frac{-\theta\phi(W_{t+1}-W)}{U'(C_t)}=-\theta\phi\frac{R_{m,t+1}(W_t-C_t)-W}{U'(C_t)}$$

因此 M_{t+1} 是 $R_{m,t+1}$ 的线性函数（以 C_t 和 W_t 为条件）。注意，为了证明 M_{t+1} 是 $R_{m,t+1}$ 的线性函数，价值函数必须只包含财富变量。如果财富的边际价值依赖于其他变量（如 $Z=$ 风险资产最近的累积损失），那么我们有 $V'(W_{t+1},Z_{t+1})$，$M_{t+1}=f(R_{m,t+1},Z_{t+1})$，单因子“标准 CAPM 模型”将不再成立。我们将在第 14 章和第 15 章考察包含财富和其他变量的价值函数。

注意，在上述多期问题中，我们假设效用函数是二次函数，这是一个限制性很强的假设，但可能是微小变化的一个合理近似（二阶 Taylor 级数展开）。为了将 $V(W_{t+1})$ 写成只是关于财富的函数，为什么需要独立同分布的假设呢？在 $t=0$ 的时刻，投资者选择 $(C_0, C_1, \cdots, C_t, \cdots)$ 和资产份额 $(\alpha_0, \alpha_1, \cdots, \alpha_t, \cdots)$ 来最大化目标函数

$$V(W_t)\equiv E_t\sum_{j=0}^{\infty}\theta^j U(C_{t+j})$$

约束条件为

$$W_{t+1}=R_{m,t+1}(W_{t+1}-C_t)\qquad R_{m,t+1}=\alpha_t' R_{t+1}\qquad \alpha' e=1$$

其中，e 是数值全为 1 的向量，$R_t=(R_1, R_2, \cdots, R_N, R_f)'$。注意，在预算约束方程中没有劳动收入。如果风险资产的收益不独立同分布（且无风险利率不为常数），那么风险资产的最优投资组合份额 $\alpha_{0t}=(\alpha_{01}, \alpha_{02}, \cdots, \alpha_{0n})$ 可以依赖于一些有助于预测未来收益的变量 $R_{t+1,t+2}$，…。例如，如果今天的低红利价格比（D/P）对未来收益具有预测功能，那么价值函数就不仅依赖于财富变量，也依赖于红利价格比，即 $V(W_t, (D/P)_t)$。这也是为了使 M_{t+1} 与 $R_{m,t+1}$ 线性相关，我们需要独立同分布和二次效用函数假设的原因。因此，SDF 跨期消费－投资组合模型无法导出标准的 CAPM 模型。在第 14 到 16 章中，我们将分析价值函数除了包括财富变量之外还包括“状态变量”的投资组合问题。

现在，我们可能会看到标准 CAPM 模型的一些缺陷。CAPM 模型假设投资者只关心孤立的投资组合问题，不关心投资组合的收益怎样与诸如收入水平等更为宽泛的经济变量发生联系。例如，在其他变量不变的情况下，如果资产是在衰退时期中进行偿付（即资产收益与产出负相关），那么人们可以预期投资者将持有那些预期收益低的资产。CCAPM 模型通过消费增长将这种负相关性在模型中进行了考虑。然而，除消费之外还可能存在影响个体效用的一些变量，这就导致了因子模型的产生（见后面）。对“跟踪者”而言后面章节将遇到的东西会足够丰富，因此我们现在转为考察价格 SDF 模型更为深层的含义及因子模型。

13.3　价格与协方差

与上述分析等价的分析是根据价格而不是收益来导出结论。我们将发现，价格依赖于预期

偿付加上“风险溢价”，而风险溢价由 M_{t+1} 与随机偿付 X_{t+1} 之间的协方差决定。

在 $P_t = E_t(MX)$ 中应用（18）并代入 $R_f^* = 1/E(M)$ 就得到

$$P_t = \frac{E_t(X_{t+1})}{R_f^*} + \text{Cov}_t(M_{t+1}, X_{t+1}) \tag{27a}$$

将 M_{t+1} 的定义式代入后就得到

$$P_t = \frac{E_t(X_{t+1})}{R_f^*} + \frac{\text{Cov}_t[\theta U'(C_{t+1}), X_{t+1}]}{U'(C_t)} \tag{27b}$$

如果投资者无任何风险规避倾向（效用函数是关于消费的线性函数）或消费是常数，那么通常的“风险中性”公式成立：

$$P_t = \frac{E_t(X_{t+1})}{R_f^*} \tag{28}$$

因此，在风险中性条件下，风险资产的价格是偿付 X_{t+1} 用无风险利率进行折现后的预期现值；在风险规避情形中，(27) 式成立，(27a) 或 (27b) 的第二项为风险调整项。例如，对正的协方差而言这一项是正的。如果在边际效用高（即消费低）时偿付 X_{t+1} 也高，那么资产的现价也就很高。对给定的 $E_t(X_{t+1})$ 而言，后者意味着（t 和 $t+1$ 之间的）低预期收益。因此，(27b) 与 (28) 是一致的。如果在边际效用高时资产的偿付 X_{t+1} 也高，那么你基本上愿意选择为资产支付一个更高的价格，这个价格高于用无风险利率对偿付进行折现后的资产现值。这是保险原理的再次应用。

上面应用的收益与价格之间的关系有时似乎很荒谬。但仔细考虑却不存在任何问题，因为 $R_{t+1}^* \equiv X_{t+1}/P_t$，所以对 $t+1$ 时刻固定的现金流 X_{t+1} 而言，今天的低价格必然意味着高收益 R_{t+1}^*。

13.4 理性定价公式与 SDF 模型

使用 SDF 模型的均衡条件，我们就能导出任何资产的理性定价公式（Rational Valuation Formula，RVF）

$$E_t(R_{t+1}^* M_{t+1}) = 1 \tag{29}$$

定义

$$R_{t+1}^* \equiv \frac{V_{t+1} + D_{t+1}}{V_t} \tag{30}$$

我们就得到

$$V_t = E_t(M_{t+1}[V_{t+1} + D_{t+1}]) \tag{31}$$

向前重复替代就得到（假设横截性条件成立）

$$V_t = E_t \sum_{j=1}^{\infty} M_{t,t+j} D_{t+j} \tag{32}$$

$$M_{t,t+j} = M_{t+1} M_{t+2} \cdots M_{t+j}$$

$M_{t,t+j}$ 是可能的时变随机折现因子。运用（18），再一次将（32）中的 $E(xy)$ 进行替换后就得到

$$P_t = \sum_{j=1}^{\infty} \frac{E_t D_{t+1}}{R_{f,t+j}^*} + \sum_{j=1}^{\infty} \text{Cov}_t(D_{t+j}, M_{t,t+j}) \tag{33}$$

这里的 $1/R_{f,t+j}^* \equiv E_t(M_{t,t+j})$ 是第 j 期的无风险利率。我们又一次看到，那些红利与边际效用之间存在负协方差（与消费之间的协方差为正）的资产具有更低的价格，从而要求一个更高的收益（对给定的预期红利流而言）。

13.5　因子模型

因子模型假定 SDF 线性地依赖于 n 个因子 $f_{i,t}$。

$$M_{t+1}=\sum_{i=1}^{n} b_{i,t}f_{i,t+1}$$

线性假设可以用一种明显的方式来导出（如关于消费的两期二次效用函数，消费在这里就是一个因子），也可以将其视为非线性关系的 Taylor 级数逼近。如果边际效用包含其他状态变量 $U_c\ (C_t,\ Z_t)$，那么 Z_t 也可能与一些因子相关。

条件模型和无条件模型

我们在上面已经看到，一般情况下系数 $\{a_t,\ b_t\}$ 可能具有时变性，对标准 CAPM 模型（单因子模型，因子为 $f_{t+1}=R_{m,t+1}$）而言，

$$M_{t+1}=a_t+b_tR^*_{m,t+1}$$

其中，$R^*_{m,t+1}$为“市场”或“财富”组合（所有资产）的总收益。应用 $E_t\ [R^*_{i,t+1}M_{t+1}]$ 就得到条件模型

$$1=E_t\ [\ (a_t+b_tR^*_{m,t+1})\ R^*_{i,t+1}]$$

取无条件期望，

$$1=E\ (a_tR^*_{i,t+1})\ +E\ [b_t\ (R^*_{m,t+1}R^*_{i,t+1})]$$

利用 $E\ (xy)\ =E\ (x)\ E\ (y)\ +\mathrm{Cov}\ (x,\ y)$，我们有

$$1\ =E\ (a_t)\ E\ (R^*_{i,t+1})\ +\mathrm{Cov}\ (a_t,\ R^*_{i,t+1})\ +E\ (b_t)\ E\ (R^*_{m,t+1}R^*_{i,t+1})\ +\mathrm{Cov}\ (b_t,\ R^*_{m,t+1}R^*_{i,t+1})$$

因此，除非协方差项等于零，否则条件 CAPM 模型成立并不意味着无条件 CAPM 模型成立，也就是说

$$1\neq E_t\ [\ \{E\ (a_t)\ +E\ (b_t)\ R^*_{m,t+1}\}\ R^*_{i,t+1}]$$

而在一般情况下协方差项不等于零（如对二次效用函数而言，a_t 和 b_t 依赖于 C_t，后者不可能与 $R^*_{i,t+1}$或 $R^*_{m,t+1}$无关）。当然，如果 a 和 b 一直为常数，那么条件模型成立确实意味着无条件模型成立，因为

$$1=E_t\ [\ (a+bR^*_{m,t+1})\ R^*_{i,t+1}]$$

意味着

$$1=E\ [\ (a+bR^*_{m,t+1})\ R^*_{i,t+1}]$$

在使用因子模型时的一个问题是，我们希望系数 a_t 和 b_t 是常数，因为欧拉方程（Euler Equation）可以用常系数 GMM 方法来进行估计

$$E\ [\ (a+bf_{t+1})\ R^*_{i,t+1}]\ =1$$

减轻这一问题影响的一个“技巧”是假设 $\{a_t,\ b_t\}$ 依赖于 t 时刻的一族可观察变量 z_t（如红利价格比），目的是为了使

$$M_{t+1}=a\ (z_t)\ +b\ (z_t)\ f_{t+1}$$

如果 a 和 b 是 z_t 的线性函数，那么

$$M_{t+1}=a_0+a_1z_t+b_0f_{t+1}+b_1\ (z_tf_{t+1})$$

我们用一个固定（常）系数的三因子模型来代替具有时变系数的单因子模型，还可以使用标准的 GMM 估计技巧。在实践中，我们可能选取好几个变量 z，从而模型包含 z 和因子 f_k的乘积项。

潜因子方面的文献都假设无法观察的过程都可以用一些因子来进行设定，而更为普遍的方法是选择较小数目的显式变量，这些变量被认为会对未来消费（如产出增长、通货膨胀率

等）产生影响。这一领域文献的另一个分支是假设f_i为那些能刻画系统风险来源的“因子模拟组合”（如 Fama - French 的三因子模型就可以用这种方式进行解释）。使用因子模拟组合的一个理由在于如下事实：在对消费（CCAPM 模型中的“正确”变量）进行测度时存在误差，所选择的因子要能对那些影响当前和未来消费的变量进行更好的测度。然而，“钓鱼”探险的范围和数据挖掘是这一领域要考虑的问题。另外，如果投资者使用那些被计量经济学家所省略的条件信息，那么对潜在 CCAPM 模型进行检验的过程中很明显充满了解释方面的困难，人们总是以“无法观察的”因子为由来拒绝条件 CCAPM 模型（Hansen 和 Richard（1987））。

如果我们将这些因子模型视为 CCAPM 模型的特殊情形，那么风险资产的超额收益由因子与资产收益之间的所有协方差来确定。他们也是共同的因子，确定资产 i 收益的任意一个因子也会对资产 j 的收益产生影响。当然，标准 CCAPM 模型可以视为因子模型，模型中的单一因子是市场投资组合的收益。

13.6 小结

* 在 SDF 模型中，用无风险利率、任意风险资产 *i* 的收益和价格表示的核心方程（对任意个体都成立）为

$$R_{ft}^{*} = \frac{1}{E_t\ (M_{t+1})}$$

$$1 = E_t\ (R_{i,t+1}^{*} M_{t+1})$$

$$P_{it} = E_t\ (X_{i,t+1} M_{t+1})$$

这里的M_{t+1}为随机折现因子。对 CCAPM 模型而言，$M_{t+1} = \theta U'\ (C_{t+1})\ /U\ (C_t)$。增加代表性投资者的假设后模型意味着上述方程可以确定均衡收益。

* 在 CCAPM 模型中，风险溢价取决于资产 i 的收益与 SDF 之间的（条件）协方差，风险溢价具有时变性。

* 在 CCAPM 模型中，如果效用函数为幂效用函数，消费增长和资产收益服从联合对数正态分布，那么资产 *i* 的风险溢价取决于可观察的消费增长与资产 *i* 收益之间的条件协方差。（实际）利率也与消费的条件方差负相关（谨慎储蓄效应），与消费的预期增长正相关。

* CCAPM 模型意味着资产价格依赖于预期红利和未来的随机折现因子$M_{t,t+j}$。

* 标准 CAPM 模型可以视为 CCAPM 模型的一种特殊情形，在标准 CAPM 模型中，SDF 与市场投资组合的收益R_m完全相关。然而，没有任何理由可以证明这一假设在一般情况下成立，尽管这种完全相关性可以视为非线性关系的一阶 Taylor 级数逼近，或在限制性假设下（如两期二次效用函数）用一种更为明显的方式来导出这种完全相关性。

* 因子模型假设边际替代率M_{t+1}线性地依赖于一组因子（我们对这些因子的度量比对消费要更准确一些）。这允许预期收益依赖于因子和资产收益之间的所有条件协方差。

* 正如 Cochrane（2001）所解释的那样，上述三个方程所表示的 CCAPM 模型并没有正态分布收益、独立同分布收益、两期投资者、二次效用函数或无劳动收入等假设。当然，为了使模型更加易于使用或更有兴趣，我们可以考虑这些假设中的一个部分。

附录：联合对数正态分布与幂效用函数

在 SDF 模型中，联合对数正态分布是一个频繁使用的简化假设。我们在这里将证明如何用联合对数正态分布来获得风险资产收益和无风险利率的一些更易于使用且更为直观的结果。

A. 联合对数正态分布

请证明：

当随机折现因子 M_{t+1} 和资产总收益 R^*_{t+1} 服从联合对数正态分布时，

$$E_t\ (M_{t+1}R^*_{t+1})\ =1 \tag{A1}$$

意味着无套利均衡条件成立：

$$E_t\ (\ln R^*_{t+1}-\ln R^*_{ft})\ +\frac{1}{2}\sigma^2_t\ (\ln R^*_{t+1})\ =-\mathrm{Cov}_t\ (\ln M_{t+1}.\ \ln R_{t+1}) \tag{A2a}$$

或

$$E_t\ (r^*_{t+1}-r^*_{ft})\ +\frac{1}{2}\sigma^2_t\ (r^*_{t+1})\ =-\mathrm{Cov}_t\ (m_{t+1},\ r^*_{t+1}) \tag{A2b}$$

小写字母表示取对数后的情形，$\sigma^2\ (r^*_{t+1})$ 为 Jensen 效应，协方差项为风险溢价。因此，任意风险资产的预期超额收益（负向地）依赖于随机折现因子与资产收益之间的协方差。

证明：

令 $\ln X\sim N\ (\mu,\ \sigma^2)$，其中，$\mu=E\ (\ln X)$，$\sigma^2$ 是“$\ln X$”的方差。因此，X 服从对数正态分布，从而如下标准结果成立

$$E\ (X)\ =e^{(\mu+\sigma^2/2)}\equiv e^{E(\ln x)+\sigma^2(\ln x)/2} \tag{A3}$$

其中，$\sigma^2\ (\ln X)$ 为“$\ln X$”的协方差。方程（A3）两边同时取对数后得

$$\ln E\ (X)\ =E\ (\ln X)\ +\sigma^2\frac{\ln X}{2} \tag{A4}$$

令 $X=\ (M_{t+1}R^*_{t+1})$，方程（A4）和（A1）两边取条件期望后就得到

$$\ln E_t\ [M_{t+1}R^*_{t+1}]\ =E_t\ [\ln\ (M_{t+1}R^*_{t+1})]\ +\frac{\sigma^2_t\ (\ln\ [M_{t+1}R^*_{t+1}])}{2} \tag{A5a}$$

因此

$$E_t\ (m_{t+1})\ +E_t\ (r^*_{t+1})\ +\frac{\sigma^2_t\ (m_{t+1})}{2}+\frac{\sigma^2_t\ (r^*_{t+1})}{2}+\mathrm{Cov}_t\ (m_{t+1},\ r^*_{t+1})\ =0 \tag{A5b}$$

当 R^* 为非随机无风险利率时，

$$E_t\ (r^*_{t+1})\ =r^*_{ft}$$

$$\sigma^2_t\ (r^*_{ft})\ =\mathrm{Cov}\ (m_{t+1},\ r^*_{ft})\ =0$$

方程（5b）是无风险利率的确定方程

$$E_t\ (m_{t+1})\ +r^*_{ft}+\frac{1}{2}\sigma^2_t\ (m_{t+1})\ =0 \tag{A6}$$

或

$$r^*_{ft}=-E_t\ (m_{t+1})\ -\frac{1}{2}\sigma^2_t\ (m_{t+1})$$

（A5）减（A6）后得到对数正态分布假设下任意资产 i 的无套利均衡条件（A2a）或（A2b）：

$$E_t\ (r^*_{t+1}-r^*_{ft})\ +\frac{1}{2}\sigma^2_t\ (r^*_{t+1})\ =-\mathrm{Cov}_t\ (m_{t+1},\ r^*_{t+1}) \tag{A2b}$$

注意，如果 m_{t+1} 和 r^*_{t+1} 具有条件同方差，从而

$$\mathrm{Var}_t\ (x)\ =E_t\ (x-E_tx)^2=\mathrm{Var}\ (x-E_tx)^2$$

那么我们省略掉方差和协方差项的下标 t（但 $E_t\ (r^*_{t+1}-r^*_{ft})$ 的下标 t 却不能省略）。将（A2）和不需要对数正态分布假设的无套利方程（参见方程（21），即如下方程）

$$E_tR^*_{t+1}-R^*_{ft}=-R^*_{ft}\mathrm{Cov}_t\ (M_{t+1},\ R^*_{t+1}) \tag{A7}$$

进行比较后就会发现，关键差异在于（A2）中多了一项（1/2）$\sigma^2\ (r^*_{t+1})$，即 Jensen 不等式项（JIT）。除非 $f\ (x)$ 是线性函数，否则在对非线性函数取数学期望时就会产生这一项，因为 $E\ [f\ (x)]\ \neq f\ [E\ (x)]$。

B. 对数正态分布加幂效用函数

现在，我们加入幂效用函数的假设并导出用无风险资产和任意风险资产的预期收益表示的方程。应用幂效用函数就得到

$$M_{t+1} = \theta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \tag{A8}$$

$$m_{t+1} = \ln\theta - \gamma\Delta c_{t+1} \tag{A9}$$

条件方差为

$$\sigma_t^2\ (m_{t+1}) = \gamma^2\sigma_t^2\ (\Delta c_{t+1}) \tag{A10}$$

将（A9）和（A10）代入（A6），对数正态分布和幂效用函数假设下的均衡无风险利率为

$$r_{ft}^* = -\ln\theta + \gamma E_t\ (\Delta c_{t+1}) - \frac{\gamma^2}{2}\sigma_t^2\ (\Delta c_{t+1}) \tag{A11}$$

这就是正文中的方程（16）。现在转为讨论风险资产的收益，根据条件协方差的定义并应用（A9），我们有

$$\begin{aligned}\mathrm{Cov}_t\ (m_{t+1},\ r_{t+1}^*) &\equiv E_t\ [\ (m_{t+1} - E_t m_{t+1})\ (r_{t+1}^* - E_t r_{t+1}^*)\ ]\\ &= -\gamma\mathrm{Cov}_t\ (\Delta c_{t+1},\ r_{t+1}^*)\end{aligned} \tag{A12}$$

（*A*12）左边无法观察的协方差现在是用消费增长和资产收益之间可观察的协方差来表示的（要用 γ 进行尺度变换，γ 为相对风险规避系数或跨期替代弹性的倒数）。考虑（A2b）

$$\begin{aligned}E_t\ (r_{t+1}^* - r_{ft}^*) + \frac{1}{2}\sigma_t^2\ (r_{t+1}^*) &= -\mathrm{Cov}_t\ (m_{t+1},\ r_{t+1}^*)\\ &= \gamma\mathrm{Cov}_t\ (\Delta c_{t+1},\ r_{t+1}^*)\end{aligned} \tag{A2b}$$

并使用如下逼近技巧

$$r_{t+1}^* \equiv \ln R_{t+1}^* \approx R_{t+1}$$

$$\sigma_t^2\ (r_{t+1}^*) \approx \sigma_t^2\ (R_{t+1})$$

$$\mathrm{Cov}_t\ (\Delta c_{t+1},\ r_{t+1}^*) \approx \mathrm{Cov}_t\ (\Delta C_{t+1}/R_{t+1})$$

我们在对数正态和幂效应函数假设下有如下逼近

$$E_t\ (R_{t+1} - R_{f,t}) + \frac{1}{2}\sigma_t^2\ (R_{t+1}) = \gamma\mathrm{Cov}_t\left(\frac{\Delta C_{t+1}}{C_t},\ R_{t+1}\right) \tag{A13}$$

第二项为 Jensen 效应，风险溢价取决于消费增长和风险资产收益之间的条件协方差。消费增长与风险资产收益之间的协方差越高，如果投资者愿意持有这些资产，那么风险资产的预期收益就越高。但是，如果我们考虑预期总收益的对数，那么我们就能消除 Jensen 不等式项，因为

$$\ln\left\{E_t\left(\frac{R_{t+1}^*}{R_{f,t}^*}\right)\right\} = \gamma\mathrm{Cov}_t\ (\Delta c_{t+1},\ R_{t+1}) \tag{A14}$$

协方差是随机折现因子模型的“国王”，因为协方差决定了风险资产或风险资产组合的条件预期收益。

第 14 章

CCAPM 模型：证据与扩展

目的

* 考察 CCAPM 模型是否与短期股票收益近似不可预测而在较长时期内却有较强可预测性的典型事实相一致。
* 考察 CCAPM 模型是否与股票收益的高平均股权风险溢价和无风险利率的低均值低波动率典型事实相一致。这两个典型事实就是所谓的股权溢价之谜和无风险利率之谜。
* 考察 CCAPM 模型欧拉方程（或一阶条件）的估计值是否给出了风险规避程度的“合理”估计。
* 分析那些包含诸如财富或过去消费等状态变量的更为一般的（如不可分离的）效用函数是否改善了随机折现因子（SDF）模型的实证绩效。

14.1　在 CCAPM 模型中资产收益是否可预测?

随机折现因子（SDF）模型与如下典型事实相一致吗：资产的短期（如日内、每日或每周）收益完全无法预测，从而资产价格服从鞅过程？记住，如果

$$P_t = E_t P_{t+1} \quad 或 \quad P_{t+1} = P_t + \varepsilon_{t+1} \tag{1}$$

P_t 是鞅。其中，$E_t \varepsilon_{t+1} = 0$，$\varepsilon_{t+1}$ 独立于 P_t。对鞅而言，（零红利）收益 $E_t P_{t+1}/P_t = 1$ 为常数，如果 ε_t 独立同分布，那么价格服从随机游走。SDF 模型的一阶条件为

$$P_t U'(C_t) = E_t[\theta U'(C_{t+1})(P_{t+1} + D_{t+1})] \tag{2}$$

因此，如果

(a) 投资者是风险中性的（即 $U(C)$ 是线性函数，从而 $U'(C)$ 为常数）；

(b) T 和 $T=1$ 之间没有任何红利支付；

(c) θ 接近于 1。

那么 SDF 模型意味着 P 是鞅过程。

CCAPM 模型与如下典型事实是一致的：日内和每日股票、债券和即期外汇价格的变化近似不可预测，因为上述假设（a）～（c）在短期内是合理的。当然，如果每日价格变化确实不可预测，那么这就使“技术分析”（如表格、蜡烛图和神经网络）这一赚钱方法（用风险和交易成本进行调整）变得无效。

较长时期

长期股票收益（如一到五年的时期）似乎呈现出一定的可预测性（尽管即将揭示的关系

在不同时期上不必是稳定的）。CCAPM 模型与典型事实相一致吗？任意风险资产收益 R 的 CCAPM 模型表示为

$$
\begin{aligned}
E_t R_{t+1}^{*} - R_{f,t}^{*} &= \frac{-\mathrm{Cov}_t\ (M_{t+1},\ R_{t+1}^{*})}{E_t\ (M_{t+1})} \\
&= -\left[\frac{\sigma_t\ (M_{t+1})}{E_t\ (M_{t+1})}\right]\sigma_t\ (R_{t+1}^{*})\ \rho_1\ (M_{t+1},\ R_{t+1}^{*})
\end{aligned}
\tag{3}
$$

其中 $R_{t+1}^{*} \equiv (1 + R_{t+1})$，其余类似。如果我们现在假设（为简化起见）效用函数是幂效用函数，消费增长服从对数正态分布，那么

$$
E_t R_{t+1}^{*} - R_{f,t}^{*} \approx [\gamma_t \sigma_t\ (\Delta c_{t+1})]\ \sigma_t\ (R_{t+1}^{*})\ \rho_t\ (M_{t+1},\ R_{t+1}^{*})
\tag{4}
$$

在将 R_{t+1}^{*} 用 R_{t+1} 替换后，上述两个方程都成立。我们没有理由认为方程（4）右边的任何一个条件矩不会随着时间而发生改变。然而，预测收益波动率的那些变量似乎并不能预测预期收益（如 Schwert（1989），ρ_T 解释也有点棘手。因此，解释均衡超额收益随着时间变化的重要竞争者要么包括时变风险规避系数 γ_t，要么包括消费的时变波动率。正如我们即将看到的那样，时变风险规避系数被广泛用于在 SDF 框架内"解释"所观察到的时变超额收益。

然而需要注意的是，我们不能将所有"收益可预测""厨房中水槽"式的时间序列回归方程与 γ_t 或 σ_t（Δc_t）的变化对应起来。我们能说的最好结果是，这些收益可预测性方程中所使用的变量（如红利价格比率、通货膨胀率、到期收益等）对 γ_t 或 σ_t（Δc_t）的变化具有"指示"作用。但这并不令人放心，或者说，这并不是很科学。

Hansen - Jagannathan 边界和 Sharpe 比

SDF 模型意味着折现因子有一个下界，我们可以将这个下界与观察到的 Sharpe 比进行比较，因此，这个下界为特殊效应函数假设下的 SDF 模型提供了一种检验方法。这种方法的"弹性"在于，我们可以在种类众多的效用函数及多个风险规避系数值基础上来计算这个边界。例如，这使得我们能观察风险规避系数的哪些可能值与给定效用函数是一致的。如果对给定的效用函数而言风险规避系数的任何值"在直觉上都不是可接受的"，那么为了进行更为深入的考察我们可以放弃这个特殊的效用函数（纯粹主义所采用的就是这种方法）。在这一节中，为方便起见，我们省略时间下标（即 $E \equiv E_t$，$M \equiv M_t$）。

运用 Cov（R_t^{*}，M）$\equiv \rho_{iM}\sigma$（R_i^{*}）σ（M）并将 $1 = E$（MR_i^{*}）代入就得到

$$
1 = E\ (MR_i^{*})\ = E\ (M)\ E\ (R_i^{*})\ + \rho_{iM}\sigma\ (R_i^{*})\ \sigma\ (M)
\tag{5}
$$

进行重新排列并运用 $E\ (M)\ = 1/R_f$，

$$
\frac{ER_i^{*} - R_f^{*}}{\sigma\ (R_i^{*})} \equiv \frac{E\ (R_i^{*e})}{\sigma\ (R_i^{*e})} = -\rho_{iM}\frac{\sigma\ (M)}{E\ (M)}
\tag{6}
$$

这里的 R_i^{*e} 为超额收益。左边的项是风险资产或风险资产组合的 Sharpe 比。对任意投资组合 i 而言，方程（6）给出了折现因子的 Hansen 和 Jagannathan（1991）边界。因为 ρ_{iM} 绝对值的最大值为 1，所以

$$
\frac{\sigma\ (M)}{E\ (M)} \geqslant \frac{|\ E\ (R_i^{*e})\ |}{\sigma\ (R_i^{*e})}
\tag{7}
$$

假设右边是可测的，方程（7）给出的是 SDF 行为的下界（也是"Hansen Jagannathan 边界"的一个简单版本，参见下文）。我们也能看出方程（7）对位于"有效前沿"上"特殊"投资组合的含义，无风险资产不存在就是这种情形——这就是图 1 中的楔形区域（它的上部等价于 CML）。我们现在能将 SDF 方法与标准 CAPM 类型的均值方差模型联系起来。

首先，所有资产收益都位于楔形区域内。将风险资产与那些在给定预期收益水平下具有最小方差的投资组合结合起来，我们就得到均值方差有效组合（当我们拥有风险资产时）。楔形区域的前沿与随机折现因子 M 完全相关，$|\ \rho_{iM}\ | = 1$。上部资产（即与 CML 等价）的收益与 M 完全负相关，从而与消费完全正相关。根据方程（6），这些资产所要求的预期收益最高。

反之，对位于楔形区域下部的那些资产（或投资组合）而言，$\rho_{iM}=+1$，因此它们的预期收益也较低，因为 R_i 和消费是完全负相关的，所以这些资产为消费变化提供了一种保险。

楔形均值方差前沿上任何两种资产收益（$R_{mv,1}$，$R_{mv,2}$）之间都是完全相关的（因为每种资产的收益都与 M 完全相关），因此

$$R_{mv,1}=R_f+\delta\ (R_{mv,2}-R_f)$$
$$M=a+bR_{mv,j}$$
$$R_{mv,j}=c+dM$$

其中，j 是均值方差有效前沿上的任何一个投资组合。最后两个方程说明，前沿上的任何收益（即 $R_{mv,j}$）包含了定价核或随机折现因子 M 的所有定价信息。最后，可以证明任意资产的预期超额收益都与相对于有效前沿上任意资产收益的 $\beta_{i,mv}$ 成比例

$$ER_i-R_f=\beta_{i,mv}\ (ER_{mv}-R_f)$$

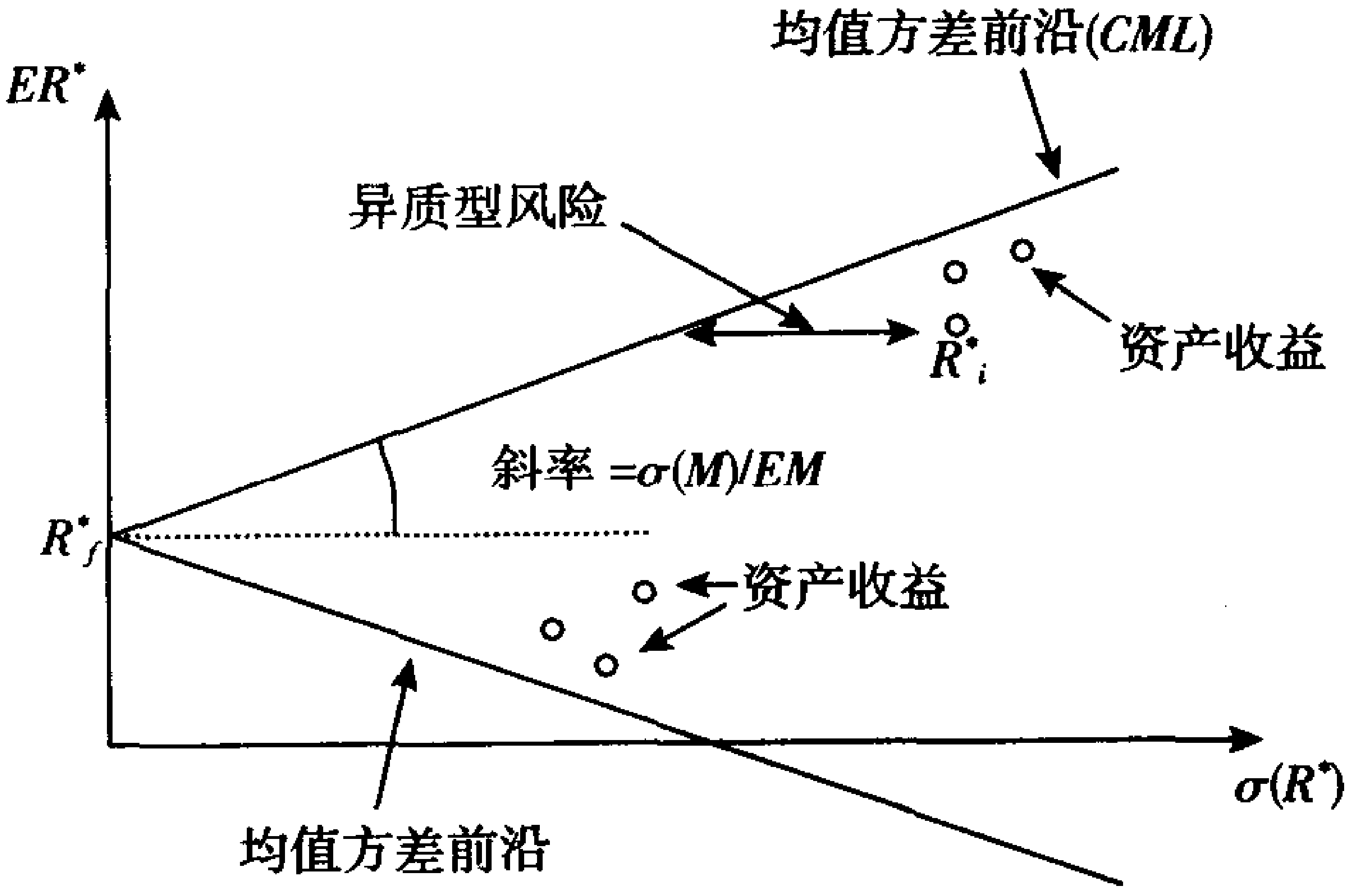


图 1　均值方差前沿（存在无风险资产）

（平均）收益 ER_i（$i=1,2,\cdots,N$）的横截面图像与 $\beta_{i,mv}$ 线性相关。因为上述方程也适用于 β 值为 1 的资产 R_{mv}，所以因子风险溢价为 $\lambda_M=ER_{mv}-R_f$。对任意资产而言，根据方程（3）我们有

$$\begin{aligned}E_tR_{i,t+1}-R_f&=\frac{-\mathrm{Cov}_t\ (M_{t+1},\ R_{i,t+1})}{E_t\ (M_{t+1})}\\&=\frac{\mathrm{Cov}_t\ (M_{t+1},\ R_{i,t+1})}{\mathrm{Var}_t\ (M_{t+1})}\left(\frac{-\mathrm{Var}_t\ (M_{t+1})}{E_t\ (M_{t+1})}\right)\\&=\beta_{i,M}\lambda_M\end{aligned}$$

其中，$\beta_{i,M}$ 为 R_i 对 M 进行回归的回归系数，而 λ_M 对任意资产而言都相同，可以将其解释为风险的市场价格。对幂效用函数而言，$M_{t+1}=\beta\ (C_{t+1}/C_t)^{-\gamma}$，可以证明（对上述方程进行 Taylor 级数展开）

$$E_tR_{i,t+1}-R_f\approx\beta_{i,\Delta C_{t+1}}\ [\gamma\mathrm{Var}_t\ (\Delta C_{t+1})]$$

因此，效用函数为幂效用函数的 SDF 模型意味着风险的市场价格正向地依赖于 γ 和消费的风险。标准 CAPM 模型（均值方差模型）和 SDF 模型之间是相互平行的，从上述方程就可以看出这一点。在 SDF 模型中 M_{t+1} 是一个普通变量，市场收益 R_m 只是一种特殊情形。（注意，在上述所有方程中我们可以用 R^* 来替换 R。）

我们现在转为讨论有效前沿（CML）内任意投资组合 Sharpe 比的含义。根据方程（6），应用 $|\rho_{iM}|=1$，我们有

$$SR_{mv}=\left|\frac{E\ (R^*_{mv})\ -R^*_f}{\sigma\ (R^*_{mv})}\right|=\frac{\sigma\ (M)}{E\ (M)}=\sigma\ (M)\ R^*_f \qquad (8)$$

因此，有效前沿上任意投资组合都有相同的 Sharpe 比，决定 Sharpe 比大小的“基础”是无风险利率和 SDF 的波动率。因为这两个变量都可能随着时间发生变动，我们也预期所获得的 Sharpe 比具有时变性。运用幂效用函数 $U'(C)=C^{-\gamma}$和对数消费增长的假设，我们有

$$\begin{aligned}\frac{\sigma(M)}{E(M)}&=\frac{\sigma\left\{\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right\}}{E\left\{\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right\}}\\&=\left[e^{\gamma^2\sigma^2(\Delta c_{t+1})}\right]^{\frac{1}{2}}\\&\approx\gamma\sigma(\sigma c)\end{aligned}\tag{9a}$$

因此，由方程（8）和（9a），如果消费增长易于波动或 γ 很大，那么所观察到的有效前沿内资产或投资组合的高 Sharpe 比与 CCAPM 模型是一致的。消费增长波动率和 γ 两者都是经济“风险”的度量指标。随着消费增长条件波动率的变化，Sharpe 比也会发生变化。对那些没有位于有效前沿（CML）上的资产 i 而言，Sharpe 比（根据（6））在幂效用函数和对数正态分布假设下应等于

$$SR_i\approx-\rho_{iM}\gamma\sigma(\Delta c)\tag{9b}$$

14.2 股权溢价之谜

Mehra 和 Prescott（1985）关于股权溢价之谜的学术论文是建立在 SDF 模型基础上的，如下假设也是需要的：

（ⅰ）关于消费的标准偏好（如幂效用函数）；

（ⅱ）投资者最大化一生的效用，效用函数只依赖于消费且具有时间可分性；

（ⅲ）资产市场完备，投资者可以通过签订保险合同来应对不确定性（如大量失业）；

（ⅳ）资产在无摩擦的市场中进行交易，因此没有任何交易成本（即佣金费用、税收等都可以忽略不计）。

可以证明，假设（ⅲ）和（ⅳ）意味着存在“代表性投资者”，从而（非耐用品和服务上的）个体消费与人均消费变化的幅度相同。注意，代表性投资者模型的一个充分条件是，所有个体都具有相同的偏好（效用函数），都拥有相同的生产可能性所有权（ownership of production possibikities），因此他们的消费相同。但是，这是一个不必要的假设，关键在于资产市场是完备的无摩擦市场（Constantinides（1982））。

我们假设标准普尔 500 指数 R_m 总是在有效前沿的上部，因此 $\rho_{iM}=-1$。在过去 50 年中（即二战后），年平均实际收益约为 9%，标准差为 16%，年平均实际无风险利率约为 1%，SR =0.5。消费增长率的标准差约为 1%。如果（9a）的 CCAPM 能“拟合”这些典型事实，那么所要求的风险规避系数为 $\gamma=50$。根据赌博试验，一般认为 γ 在 3 ~ 10 的范围内取值。这就是著名的股权溢价之谜。

你是风险规避的吗？

无论是通过内省的方法还是用试验经济学的方法来确定风险规避程度都是非常困难的，试验经济学方法是在试验条件下让人们陈述“赌博”的范围。考察相对风险规避系数 γ 合理的不同取值的一种方法是，考虑各种赌博的确定性等价数量（正如第一章所勾画的那样）。假设我们拥有幂效用函数，初始年消费（或收入）水平为 $C_0=50\ 000$ 美元。你愿意每年支付多少来避免一个（在你余生中）每年给你增加或减少 5 美元公平的赌博呢？就幂效应函数而言，答案是满足如下条件的 z 美元

$$(C_0-z)^{1-\gamma}=\frac{1}{2}(C_0+y)^{1-\gamma}+\frac{1}{2}(C_0-y)^{1-\gamma}$$

由此得到

$$z = C_0\left\{1 - \left[\frac{1}{2}\left(1 + \frac{y}{C_0}\right)^{1-\gamma} + \frac{1}{2}\left(1 - \frac{y}{C_0}\right)^{1-\gamma}\right]^{\frac{1}{1-\gamma}}\right\}$$

可以证明，对所有的赌博而言它可以由下式来进行逼近

$$\frac{z}{y} = \frac{\gamma}{2}\left(\frac{y}{C_0}\right)$$

$$\left(\frac{\text{避免赌博的数量}}{\text{赌博的规模}}\right) = \frac{\gamma}{2}\left(\frac{\text{赌博的规模}}{C_0}\right)$$

其中 γ 是局部曲率，$\gamma = -CU''(C)/U'(C)$。因此，就小赌博而言，幂效应函数是合理的，因为，如果赌博的规模 y 相对于初始消费水平 C_0 而言很小，那么为了避免赌博你愿意支付的数量也很小（相对于赌博的规模而言）。

表 1　为避免一个公平赌博你愿意支付的美元数量

赌博数量（美元）	风险规避程度 γ				
	2	10	50	100	150
10	0.002	0.01	0.05	0.1	0.25
100	0.2	1.0	5.0	9.9	24.0
1000	20	99	435	655	863
10 000	2000	6920	9430	9718	9888
20 000	8000	17 600	19 573	19 789	19 916

注：

（ⅰ）初始消费（收入）水平为每年 50 000 美元，我们假设效用函数为幂效用函数（即相对风险规避系数 γ 为常数）。

（ⅱ）在本书的网页上可以找到 GAUSS 程序。

表 1 对此进行了说明，我们使用关于 z 的上述方程发现，即使 $\gamma = 50$，为了避免在你 50 000美元初始消费（财富）基础上增加或减少 100 美元的一个公平赌博，你愿意支付一个相对较小的数量 5 美元。这是因为，对金额很小的赌博而言，效应函数的局部曲率接近于线性（即风险中性）。然而，如果赌博是增加或减少 10 000 美元（或更多），那么当 $\gamma = 50$ 时，你愿意支付近 10 000 美元（即 9430 美元）来避免这一赌博。你对金额很高的赌博具有更强的风险规避性（这就是 Rabin（2000）矛盾的基础）。注意，即使你的初始消费（收入）水平为 300 000 美元，那么当时，你愿意支付多达6092 美元来避免增减 10 000 美元的公平赌博。与初始消费水平为 $C_0 = 50\ 000$ 美元的消费者相比，你愿意支付更少的金额，因为效用函数在更高的初始消费水平上变得“更加扁平”（边际效用递减）。表 1 中用汽车被盗或你的房屋被烧的概率所计算的数据可以用来与实际支付的保险费用进行比较，为了使数据与保费一致，我们要对 γ 进行调整。现有研究发现，人们为规避风险所愿意支付的金额比风险的精算费用高得多，这意味着 γ 的值要更低一些。因此，对我们来说，为什么人们对股票风险如此厌恶仍然是一个谜。

上述结果并不完全依赖于效用函数的函数形式。只要效用函数是财富（消费）的递增凹函数，预期效用对小额赌博而言就是合理的，但对大额赌博而言就显得有些荒谬。为了用一种稍微不同的方式来观察这一事实，我们考虑如下建议。假设我们知道，蒙帝·卡西罗先生总是拒绝损失 10 美元或获得 11 美元的 50:50 赌博，他是一个追求预期效用最大化的风险规避者（即效用函数是财富的递增凹函数）。现在，我们给蒙帝·卡西罗先生提供一个损失 100 美元或获得美元的 50:50 赌博。为了使他参与这个赌博，我们所支付的最大值为多少呢？当然，“赢” 2000 美元似乎是合理的。因果内省专家肯定建议蒙帝·卡西罗先生接受那些赢 20 000 美

元、100 万美元或 10 亿美元的赌博（记住他只能损失 100 美元）。事实上，作为一个预期效用最大化的追求者，如果他拒绝损失 10 美元获得 11 美元的 50∶50 赌博，那么他将始终拒绝损失 100 美元获得任意数量收益的 50∶50 赌博。这就是 Rabin 的自相矛盾（Rabin（1998，2000），Rabin 和 Thaler（2001）），结果似乎有些荒谬，这使得人们在财富水平是考虑赌博唯一标准的基础上使用预期效用时感到非常的心神不安。

对相同类型的重复赌博进行考虑后这一结论仍然成立。如果有人拒绝损失 X 获得 Y 的公平赌博，那么预期效用意味着这个人将拒绝参与这些相同赌博的绝大多数（Samuelson（1963））。在对不同初始财富水平进行考虑后，我们也不能（定性地）拒绝 Rabin 矛盾。例如，如果追求预期效用最大化的“富人”拒绝损失 10 000 美元获得 11 000 美元的 50∶50 赌博，那么她也将拒绝损失 100 000 获得 ∞ 的 50∶50 的赌博。上述关于风险规避预期效用最大化行为描述的观念困难导致了一些备选的效用函数出现，其中效用依赖于财富的变化，损失比收益更加“痛苦”（如损失规避或失望规避），个体孤立地（使用范围很窄）考虑收益和损失。我们在后面几章中将对此进行讨论。

从实验室获得的“试验”证据存在一个问题，个体选择违反了预期效用公理（我们在后面还要对这一问题进行讨论），远离人们通常经验的“赌博”很难进行评价。（这也适用于如下情形：个体被问到愿意支付多少金钱来减少诸如汽车事件、地铁爆炸或飞船出事等特殊类型风险发生的概率。）

如果我们接受预期效用公理，那么大的风险规避系数值并不总是与内省变化一致，这将有助于解决股权溢价之谜。但是，正如我们即将看到的那样，大的 γ 会引出其他问题，即无风险利率的低水平值和低波动性（正如在几乎所有的发展经济学中都能观察到的那样）。相对于财富赌博而言，大 γ 的一种替代方法是假设投资者也关心消费之外的其他变量，下面将进一步讨论这种情形。

最后，注意标准普尔 500 指数可能没有位于所有资产的有效前沿上，因为它只包含可能投资的一个子集，数据表明标准普尔 500 指数与总消费之间的相关系数为 0.2（而不是“理论上”的 1），因此 $\rho_{im}=0.2$。但是，运用（9b）（效用函数为幂效用函数），这意味着为了“拟合事实”，$\gamma=250$。这使得股权溢价之谜更难解释。

使用个体消费数据

如果 *CCAPM* 模型适用于每一个投资者，那么代表性投资者的假设允许我们使用加总消费数据来对模型进行检验。但是，美国市民中只有很小部分持有股票，而确定风险溢价的是那些活动积极的交易者的消费增长与资产收益之间的相关性。根据经验分析，活动积极交易者的消费增长与股票收益之间的相关性比加总消费与股票收益之间的相关性高，但是，它仍然没有高到足以解决股权溢价之谜的程度（Mankiws 和 Zeldes（1991））。事实上，借助个体（非耐用品）消费并不能解决股权溢价之谜，因为

（*a*）尽管个体消费对异质性冲击更加敏感，但人们很难相信它的变化幅度约为每年 50%（即 $\sigma(\Delta c_{i,t+1})=50\%$）；

（*b*）个体消费增长与市场收益的相关性比加总消费增长与市场收益的相关性更低（$\rho_{iM}=0.2$），这使得（9b）更难成立。

无风险利率之谜：$\gamma=50$ 可接受吗?

现在，所谓的“股权溢价之谜”实际上应视为两个谜。第一个谜在于如下事实：所观察到的（实际）每年 8% 的股权溢价实在是太高，以至于“标准偏好”模型无法解释，除非风险规避系数高得让人“无法接受”。第二个谜是无风险利率之谜（*Weil*（1992））：当风险规避系数大的时候，与根据数据研究的结果（国债的平均实际收益为每年 1%）相比，标准 *CCAPM* 模型给出的均衡无风险利率要更高一些，波动率程度也要更大一些。因此，为了使模

型能解释所观察到的股权溢价，即使高风险规避系数是可以接受的，模型也不能拟合与无风险利率运动有关的事实。为了看清这一点，注意幂效用函数和对数正态分布假设及 $\theta=0.99$：

$$\ln R_{ft}^{*}\approx R_{ft}=\ln\theta+\gamma E_{t}\ (\Delta c_{t+1})\ -\frac{\gamma^{2}}{2}\sigma_{t}^{2}\ (\Delta c_{t+1}) \tag{10}$$

其中，R_{ft}为无风险利率（用十进制表示）。$\gamma=50$ 的 CCAPM 模型对无风险利率意味着什么？（我们忽略最后一项，因为在连续时间版本的模型中没有最后一项，当很大时，最后一项是一种不合理的近似）。当 $\gamma=50$ 时，

$$R_{ft}=-0.01+50\ (0.01)\ =0.49\ (49\%p.a.) \tag{11}$$

年消费增长率近似为独立同分布序列，因此，我们应用了无条件期望，$E\ (\Delta c)\ =0.01$，$\sigma\ (\Delta c)=0.01$。这里的实际利率，由 CCAPM 进行预测（幂效用函数和对数正态分布假设），大约为每年 50%，而许多发达国家的数据表明平均实际利率约为每年 1%，两者明显不一致（如果模型包含最后一项，那么模型预测的实际利率为 36.5%）。

另外，如果 $\gamma=50$，那么当消费增长比平均水平高或低 1% 时，模型预测的实际利率将增加或减少 50%（例如，如下情形可能发生：他人在向你借钱时，你为他支付年实际利率 50%）。因此，$\gamma=50$ 意味着，对消费跨期替代的厌恶程度是如此的高，以至于利率必须出现较大的增加才能诱使人们在今天进行额外的储蓄（因此，明天的消费会更高一些）。

Cochrane（2001）用一个漂亮的例子来说明 $\gamma=50$ 对实际利率的含义。如果你每年的收入为 50 000 美元，并将收入的 5%（2500 美元）用于年度休假，那么只有当利率上升到每年

$$R_{ft}\approx\ (52\ 500/47\ 500)^{50}-1=14\ 800\% \tag{12}$$

你才会自愿放弃这一年度休假。

事实上，你可能打算将休假延迟到比这一利率更低的某个时期！因此，当效用函数为时间可分离效用函数、对数效用函数或指数效用函数时，股权溢价之谜就会产生，因为 CCAPM 模型不能同时解释所观察到的风险资产的高平均收益（和 Sharpe 比）和值小且波动性低的实际利率。

14.3　CCAPM 模型欧拉方程的检验

横截面证据

可以用横截面数据或时间序列数据来检验 CCAPM 模型。CCAPM 的横截面检验是基于如下方程：

$$ER_{i}=a_{0}+a_{1}\beta_{c,i} \tag{13}$$

其中，

$$a_{0}=\ (1-EM)\ /EM$$

$$a_{1}=a\theta\mathrm{Cov}\ (R_{m},\ \Delta c)$$

$$\beta_{c,i}=\mathrm{Cov}\ (R_{i},\ \Delta c)\ /\mathrm{Var}\ (\Delta c)$$

方程（13）是 CCAPM 模型的证券市场线（SML）。平均收益是 $\beta_{c,i}$的线性函数。在横截面数据中，EM 和 Cov（R_m，Δc）都是常数，因此对所有股票 i 而言 a_1 都相同。对任意股票或投资组合 i 而言，我们使用 ER_i 的样本均值和 β_{ci}的样本估计值。Mankiw 和 Shapiro（1986）使用 1959—1982 年期间美国 464 家公司的横截面数据对 CAPM 模型和 CCAPM 模型进行了检验。他们发现，标准 CAPM 模型的检验效果明显比 CCAPM 模型好，将 $\bar{R}_i$ 对 β_{mi}和 β_{ci}进行回归时，前者在统计上是显著的，而后者在统计上却不显著。Breedem、Gibbons 和 Litzenberger（1989）利用产业和债券投资组合数据发现了类似的结果，而 Cochrane（1996）使用按规模分类的投资组合收益横截面数据发现 CCAPM 模型比标准 CAPM 模型的效果要更差一些。然而，值得记住的

是，CAPM 模型是在市场收益给定的情形下对资产进行“定价”，CAPM 模型并不对资产收益进行解释，而 CCAPM 则要对市场收益进行解释，因为 CCAPM 模型假设市场收益取决于消费增长的波动率。换句话说，CCAPM 模型或 SDF 模型试图根据经济风险的潜在来源（如消费）来解释所有的资产收益，而不是根据其他资产的收益来进行解释。

时间序列证据

在幂效用函数假设下，CCAPM 模型中风险资产和无风险资产所对应的欧拉方程（一阶条件，FOC）分别为

$$e_{i,t+1} \equiv \theta E_t\left\{\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(R^*_{i,t+1} - R^*_{ft}\right)\right\} = 0 \tag{14a}$$

$$e_{f,t+1} \equiv \theta E_t\left\{\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R^*_{ft}\right\} = 1 \tag{14b}$$

运用迭代期望定理得出无条件期望之后，一阶条件最简单的检验方法是，针对 γ 的不同取值，取 e_i 和 e_j（及其标准差）的（时间）样本平均值。Kocherlakota（1996）报告的结果是，对加总的股票指数而言，要使一阶样本矩在统计上为零，γ 的值必须大于等于 8.5（假定 $\theta = 0.99$），这就是股权溢价之谜的“股权含义”。不幸的是，当 $\gamma = 8.5$（且 $\theta = 0.99$）时，e_f 在统计上与 1 不同，因此代表性消费者将未来消费转移到当前就能获得收益（储蓄增加），这就是无风险利率之谜。

在非线性时间序列回归中使用条件期望可以对方程（14）进行直接检验，检验结果给出了 γ 的估计值（经常要对 θ 施加限制）。如果使用风险资产的收益，那么一阶条件为

$$\left\{\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\theta R^*_{i,t+1} - 1\right\} = \varepsilon_{i,t+1} \tag{15a}$$

因为（15a）对所有风险资产（或组合）都成立，所以它是一组交叉方程限制条件，参数（θ，a）在所有方程中出现。这个模型对宽范围内的各种资产而言表现不是很好，这些资产包含在投资组合中（例如，只包括美国股票的投资组合或包括美国股票和债券的投资组合，或由不同国家股票和债券构成的投资组合，参见 inter alia，Hodrick（1987），Cumby（1990），Smith（1993））。Smith（1993）的研究使用的是“国际”一篮子资产，他发现方程（15a）中的参数限制条件一般是成立的，但参数本身关于时间不是常数。

我们也可以将方程（15a）应用到加总的股票市场收益 R_m 上，注意（15a）对任意收益期 $t+j$ 都成立，因此，对 $j=1$，2，…而言我们有

$$\left\{\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R^*_{m,t+1} - 1\right\} = \varepsilon_{i,t+j} \tag{15b}$$

方程（15b）可以用时间序列数据来进行估计，因为它对 $j=1$，2，…的所有时期都成立，我们再一次有参数“相同”的一系列方程。$j=1$，2，…时的方程（15b）与 Fama 和 French（1988b）使用不同时期收益的回归方程相似，唯一的不同在于，我们在这里隐含地嵌入了时变预期收益。Flood、Hodrick 和 Kaplan（1986）发现，从统计效果来看，随着时期的延长，CCAPM 模型的绩效变得更糟。因此，他们的发现与（这一版本的）CCAPM 相反。整体而言，根据对 γ 和跨方程限制条件的正式统计检验结果，具有幂效用函数的条件 CAPM 模型是资产收益时间序列行为的糟糕表示。

Hansen - Jagannathan 边界

对标准普尔 500 而言，Hansen - Jagannathan（H - J）边界意味着

$$\frac{\sigma(M)}{E(M)} \geqslant \frac{|E(R_i^{*e})|}{\sigma(R_i^{*e})} \approx 0.5$$

取 $E(M) = 1/R_f^* \approx 0.99$，上式意味着每年的 $\sigma(M) \approx \gamma\sigma(\Delta c) > 50\%$，而每年的 $\sigma(\Delta c)$

$\approx 1\%$，这意味着 $\gamma = 50$。正如上面所注意到的那样，如果我们承认总量消费增长与标准普尔 500 收益之间的相关性为 0.2，那么这意味着 $\gamma = 250$。

价值函数

股权溢价之谜与实际消费的均匀性或无波动性有关。你有可能被诱导，从而认为效用函数依赖于财富而不是消费，在股票市场投资给定的情况下，毫无疑问“财富组合”比消费有更大的波动率。在上一章中，我们注意到价值函数 $V(W)$ 是跨期模型一阶条件的关键内容，H－J 边界变为

$$\frac{ER - R_f}{\sigma(R)} \leqslant \frac{-WC_{ww}}{V_\omega}\sigma(\Delta\omega)$$

对持有市场投资组合的投资者而言，每年的 $\sigma(\Delta\omega) = \sigma(R_m) \approx 16\%$。如果我们希望解释 0.5 的市场夏普比，那么价值函数关于财富的局部风险规避系数的下界为

$$\frac{-WV_{ww}}{V_\omega} = \frac{0.5}{0.16} = 3.125$$

这似乎比 $\gamma = 50$ 或上面所讨论的其他结果更合理一些。然而，正如我们在第 15 章将看到的那样，如果收益独立同分布，V 只依赖于财富（例如，无劳动“收入”），那么在跨期消费模型中，消费与财富成比例，因此 $\sigma(\Delta c) = \sigma(\Delta w)$。因此，通过股市财富的高波动率和价值函数对财富的低（局部）风险规避系数来解释高夏普比，所需要的消费增长波动率为每年 16%，数据显然不支持这一结论。因此，用财富来代替消费并不能解决股权溢价之谜，这意味着实际消费的波动率不合理，从而股权溢价之谜再次出现。

对 H－J 边界进行扩展之后，可以同时用几种资产来对 CCAPM 模型进行检验。当我们拥有几种资产时，H－J 边界可用来对 SDF 施加限制，特别地，也可以对可接受的风险规避系数（对一个特殊的效用函数而言）施加限制。几种资产的 *H－J* 边界表示如下（参见附录）：

$$\sigma^2(M) \geqslant [P - (EM)(EX)]'\Sigma^{-1}[P - (EM)(EX)] \tag{16}$$

其中，P 为 $N \times 1$ 资产价格向量，X 为 $N \times 1$ 资产偿付向量，Σ 为偿付的 $N \times N$ 方差协方差矩阵。向量 P 除以初始价格后变为值全为 1 的向量 $e1$，将 X 用总资产收益向量 R^* 替换后，

$$\sigma^2(M) \geqslant [e1 - (EM)(ER^*)]'\Sigma_R^{-1}[e1 - (EM)(ER^*)] \tag{17}$$

很容易看出，对单个风险资产而言，(17) 就变为 (7)（见附录）。为了说明 H－J 边界如何解释风险规避系数可能取值方面的信息，在收益的经验方差协方差给定的条件下，我们进行如下操作。假设我们拥有股票指数实际收益和实际无风险利率的时间序列数据，那么 ER^* 为 (2×1) 向量。

如果我们用 ER^* 来替换样本平均收益 $\bar{R}^*$ 并对 (17) 右边收益的方差协方差矩阵 Σ_R 进行替换，我们就得到一个与 EM 平方有关的指标（即右边 $= a_0 + a_1E(M) + a_2[EM]^2 = f(EM, (EM)^2)$），其中 a_i 取决于估计值 $\bar{R}^*$ 和 Σ_R）。我们现在来计算

$$\sigma(M) = [a_0 + a_1E(M) + a_2(EM)^2]^{1/2}$$

在 EM 可能取值处画出 $\sigma(M)$ 关于 EM 的图像，如图 2 中的 U 形抛物线 $A - A'$ 所示。它给出了 $\sigma(M)$ 与 EM 组合的容许下界。

现在，我们来看看 H－J 下界对依赖于消费的特殊效用函数和风险规避系数的合理取值而言是否成立。首先，考虑时间可分离的幂效用函数（该效用函数具有常相对风险规避系数 γ），随机折现因子为 $M_{t+1} = \theta(C_{t+1}/C_t)^{-\gamma}$。假设 $\theta = 0.97$ 是已知的，年度 θ 值意味着实际贴现率为 3%。对任意给定的 γ，可以运用消费增长数据来计算 $\bar{M}$ 和标准差 $\sigma(M)$。因此，我们有 $\bar{M} = f_1(\gamma)$ 和 $\sigma(M) = \frac{\Sigma(M_t - \bar{M})}{T-1} = f_2(\gamma)$。我们能在 $\{\sigma(M), E(M)\}$ 空间画出与可能 γ 值相对应的点（图 2 中 $B - B'$ 曲线，用“三角形”来表示）。曲线 $B - B'$ 给出的是不同 γ 值所对应的 $\bar{M}$ 与 $\sigma(M)$ 组合（给定的是 TS－CRRA 类型的幂效用函数）。满足边界的 $B - B'$

曲线所对应的γ值是U形曲线$A-A'$内的点。例如，考虑德国加总的年度（实际）股票收益和（实际）无风险利率两种资产相应于三种效用函数的结果（Cuthebertson 和 Hyde（2004））。对熟悉的时间可分离幂效用函数（TS - CRRA）而言，图2表明位于$H-J$下界$A-A'$曲线内的$B-B'$曲线所对应的γ值为31到34，这些都是非常高的风险规避系数值。

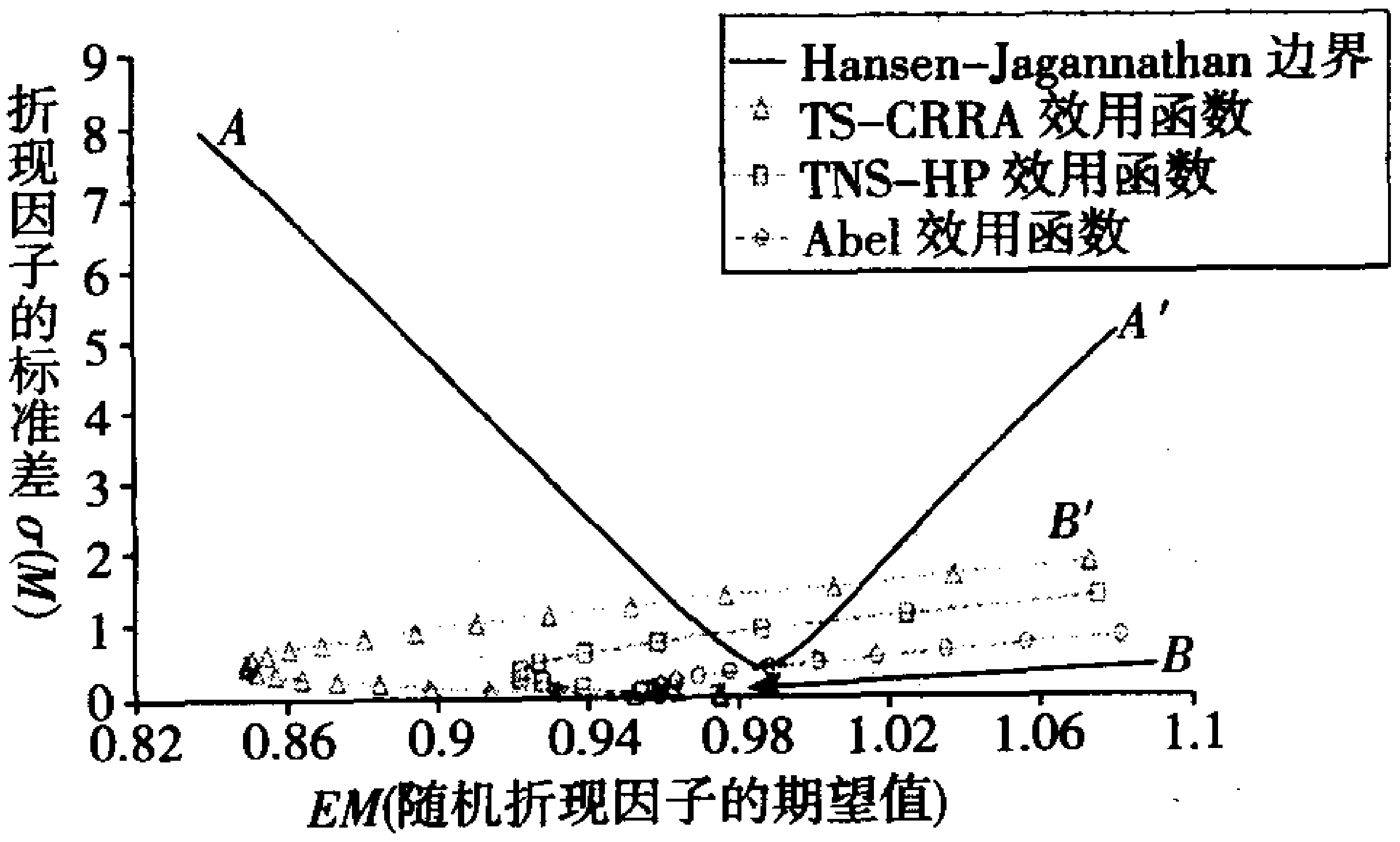


图2 Hansen - Jagannathan **边界（德国数据）**

图2也包括与两种时间不可分离的习惯持续型（habit persistence）效用函数相对应的结果，记为“TNS - HP”和“Abel”（Abel（1990））。对这些效用函数而言，人们也发现在两种偏好假设下，位于“下界”（即$A-A'$曲线）内和上的$B-B'$曲线所对应的γ值大于8小于10，这些γ值仍然太高。对美国和英国的几个研究都发现，就幂效用函数而言，容许区域内的γ值都很大，落在10～25的区域内。即使用其他效用函数如“习惯持续”型效用函数来进行替代，“合理的”γ边界（如2到4）经常被违背——参见Engsted（1998）。因此，对合理的风险规避系数和各种效用函数而言，CCAPM模型不会位于H - J边界内，这些效用函数依赖于消费增长或习惯持续型消费（Burnside（1994），Engsted（1998），Hansen - Jagannathan（1991））。这进一步说明了股权溢价之谜的存在，也进一步说明CCAPM模型的变体会出现失灵情形。

可以用正规方法来对上述结果进行检验，与H - J相关的分布理论（包括考虑买空约束的情形）、竞买/竞卖价差和其他市场摩擦等都可以在Cochrane和Hansen（1992）、Hansen、Heaton和Luttmer（1995）中找到，Cochrane（2001）也进行了非常精彩的叙述。

14.4 SDF模型的扩展

标准（幂效用函数）SDF模型不能同时解释高的股权溢价之谜和低的无风险利率之谜。绕过困境的可能途径如下：

（a）幂效用函数已经很好，但我们需要引入不可分离性，使消费的边际效用（时刻）依赖于其他状态变量（如财富、过去的消费等）；

（b）基础SDF模型是正确的，效用函数只依赖于消费，但幂效用函数是错误的。因此，我们需要其他形式的效用函数。

（c）美国股票在过去50年的收益在很大程度是上由于“好运气”，而不是对风险的偿付。这是一种“比索问题”（Peso problem），因为我们一致认为，历史数据不能反映“真实的”事前风险，正如投资者在1950—1999年期间所观察到的那样。

（d）用总量的人均消费指标来对个体消费进行度量的结果很糟糕。

幸运及夏普比的错误测度

假设我们所拥有的数据集合是给定的，股权溢价的估计值有较大的误差。取标准普尔500指数的年超额收益为8%，$\sigma = 16\%$，并用独立同分布来进行合理（“大概”（ball - park））逼近，预期（平均）收益$\bar{R}$的标准差为$\sigma/\sqrt{T}$。因此，就50年的数据而言，预期超额收益两倍标准差的置信带为8% ±2（2.3%），也就是说，大约为3.5%到12.5%，这是一个相当宽的区域。对20年的数据而言，预期收益的标准差为3.6%（$=16/\sqrt{20}$），因此当平均收益为8%时，两倍标准差置信带的下边界接近于0。然而，预期收益测度方面相对较大的不确定性可能使每一条路都滑向股权溢价之谜，因为我们要么将上限要么将下限视为“真实的”平均收益。

也许美国数据不具有典型代表性，而美国在过去50年仅仅是因为运气好而已。在1945年之后，美国经历过几场大战（例如，1945—1989年的冷战，包括对柏林的封锁，1950年的朝鲜战争，1962年的古巴导弹危机，20世纪70年代的越南战争，1992年短暂的海湾战争，2001—2003年的阿富汗战争和伊拉克战争）。没有任何一场战争导致核武器的使用，拥有“炸弹”的其他国家如俄罗斯、中国、以色列、巴基斯坦和印度都没有使用过核武器。在这些时期中美国也没有出现过严重的自然灾难和全国范围内的政治骚动（如种族暴动）。自独立以来，美国就没有任何外国侵略势力，美国本土只有两次严重的恐怖主义，一是1995年的俄克拉何马炸弹事件，因为它Timothy McVeigh被起诉，二是2001年9月11日对纽约世贸大楼双塔的攻击。

上述“幸运”场景在其他国家却不会出现。1900年，当我们以一个“世界投资者”的视角来对股权溢价进行度量时，也许我们也应该将德国、日本、中国、俄罗斯及其他卫星国包括进来，所有这些国家都经历了大的战争或暴动，股票价格从来就没有复原过。在二战后，南美和远东新兴工业化经济的收益也应该包括进来。因此，美国数据很有可能是由幸存偏好所导致的。不幸的是，在一个很长的时期内，“其他市场”高质量的连续数据也经常无法获得。

Jorion和Goetzmann（1999）试图用多个国家的长时间序列数据（1920—1996）来对预期收益和收益的标准差进行度量。尽管这种方法很难进行一般化，他们还是发现美国的实际收益率每年为4.3%，而所有其他市场（一些市场在一些时期中暂停交易）的收益仅为0.8%。然而，持有世界指数可以获得的实际收益约为每年4.3%，因为世界指数受美国主导。Dimson、Marsh和Staunton（2002）运用1900—2000年期间16个国家的历史数据重新进行了计算，结果发现美国每年的股权风险溢价为5.8%，英国为4.8%（几何平均），而16个国家作为一个整体的股权风险溢价为5%——因此，在这一时期中，美国并不是特别地不具有典型性，幸存偏好似乎相对较小。Dimson等人发现的相对较小的美国股权溢价值是因为指数构造存在细微差别，且使用了较长时期的数据。

如果我们假设“真实的”预期超额收益低到每年的3%（而不是每年的8%左右），夏普比为0.2（$=3/16$），那么CCAPM模型（效用函数为幂效用函数）和给定的$\sigma(\Delta c)=0.01$仍然需要20左右的风险规避系数（如果我们考虑$\rho_{iM}=0.2$的相关性，那么我们需要100的风险规避系数）。因此，即使美国在50年中每年8%的历史平均超额收益是一个“比索问题”（好运气），我们仍然需要超过20的风险规避系数。但我们从数据中却不能发现具有幂效用函数的CCAPM模型所预测的大且波动厉害的利率。因此，无风险利率之谜仍然存在。

问题的困难在于，为了“解释”利率的低波动性，我们需要一个模型，模型中SDF的均值$R_f^* = 1/EM$随着时间不会有太大的变化。但是，为了解释所观察到的夏普比（参见方程(7)），我们需要大的$\sigma(M)$。实际上，我们缺少一个自由度，因为单状态变量（消费）模型不能解释互相矛盾的两个事实。在效用函数中，我们需要能与随机折现因子一起变动的其他状态变量。闲暇、国外消费、历史消费水平和金融财富等都是其他状态变量的几个备用选择。

另外，我们要求模型能解释长期可预测性这一典型事实。比如，低价格红利比之后紧接着就是高收益。此外，衰退时期中股票价格较低一些，从经验来看未来收益必然很高，因此夏普

比的变化是逆周期的（即夏普比在衰退时期中高，在繁荣时期中低），我们所选择的模型需要解释这些典型事实。

不可分离性

如果边际效用依赖于其他变量 Z，即 U_c（C，Z），那么预期收益将依赖于 Z_t 和 $R_{i,t}$ 之间的协方差，也依赖于消费与 $R_{i,t}$ 之间的协方差。然而，注意如下事实：如果 U（C，Z）具有可分性，即 $U(C, Z) = f(C) + g(Z)$，那么这种情形将不会发生。需要考虑的各种不可分离性形式有：

（ⅰ）在效用函数中考虑闲暇或财富等状态变量；

（ⅱ）习惯持续性效应（因此消费的历史值和当前值都会出现在效用函数中）；

（ⅲ）所谓的非状态可分效用函数，某个状态（“热”）下消费的边际效用受其他状态下（“冷”）消费水平的影响（Epstein 和 Zin（1989））。

效用函数中的财富

Bakshi 和 Chen（1996）给出了一个跨期模型，效用函数也依赖于社会财富指数 $S_t = W_t / V_t$，其中 W_t = 个体财富，V_t = 社会财富指数。他们称该模型为“资本主义精神”模型，因为效用不仅依赖于消费，还独立地依赖于你相对于他人的“财富阶层”（“跟上琼斯”的一种类型）。他们所使用（不可分离）效用函数的形式为

$$U(C_t, W_t, V_t) = \frac{C_t^{1-\gamma}}{(1-\gamma)}\left(\frac{W_t}{V_t}\right)^{-\lambda} \tag{18}$$

$\lambda = 0$ 就简化为标准形式。跨期问题中预期收益的解为

$$ER_i - R_f = \gamma\sigma_{ic} + \lambda\sigma_{iw} - \lambda\sigma_{iv} \tag{19}$$

因为对社会财富指数 V_t 进行合适测度是困难的，所以欧拉方程的检验假设 V 为常数，从而（19）中最后一项消失。一阶条件为

$$E_t\{M_{t+1}, R_{i,t+1}\} = 1 \tag{20}$$

其中

$$M_{t+1} = \theta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R_{W,t+1}^{-\gamma}\left[1 + \frac{\lambda}{\gamma - 1}\left(\frac{C_{t+1}}{C_t}\right)\right]$$

财富组合取为总的股票市场指数，R_i 是以不同规模为基础的投资组合的收益。对不同投资组合而言，γ 的估计值落在 2.3～3.08 的范围内（λ 的取值范围为 0.75～1.27）。在检验过程中将两种风险投资组合同时包含进来后，过度识别限制不能遭到拒绝（参见 Bakshi 和 Chen（1996），表 3）。另外（对 $\lambda = 2$ 而言），当相对风险规避系数 γ 位于 6 和 8 之间时，模型并没有违反 H－J 边界。

显然，R_W 和 R_i 之间的高协方差使得模型给出的 γ 估计值很小，但仍能解释 R_i 的“波动”运动。另外，“可接受”的 λ 值存在一些偏差，无法得到实际数据的支撑（除 $\lambda > 0$ 外）。模型解决了股权溢价之谜中低 γ 部分，但似乎没法对无风险利率的欧拉方程进行检验，因此我们并不知道谜的另外一个部分。

将“状态”和“时间”分离：Epstein－Zin 模型

在标准幂效用函数中，“凹参数” γ 起着双重作用。一方面，递增的 γ 使得个体希望在不同时点匀滑消费——他们不喜欢消费增长出现太大的波动，希望消费路径匀滑。另一方面，较高的 γ 也使个体希望在不同状态下都能进行消费——他们不喜欢风险。参数 γ 既等于相对风险规避系数，又等于跨期替代弹性。

Epstein 和 Zin（1989，1991）给出了一种广义预期效用（GEU）函数，$U_t = U[C_t, E_t(U_{t+1})]$，当前消费与未来效用之间替代弹性为常数的效用函数是它的一种特殊情形。

$$U_t = [(1-\delta)\ C_t^{1-1/\psi} + \delta E_t\ \{U_{t+1}^{1-\gamma}\}^{1/\lambda}]^{1-(1-1/\psi)} \tag{21}$$

其中，$\lambda = (1-\gamma) / (1-1/\psi)$。风险规避程度由 γ 来决定，而跨期替代弹性由 ψ 确定，因此两者可以不同。从原理上讲，一个参数 γ 可以用于解释股权溢价，另一个参数 ψ 可以用于解释值小且波动率性弱的无风险利率。如果 $\gamma = 1/\psi$，我们就得到了标准的幂效用函数。实际值（而不是名义值）意义下的预算约束（无劳动收入）为 $W_{t+1} = (1+R_{m,t+1})\ (W_t - C_t)$，其中，$R_{m,t+1} = \sum_{j=1}^{n} \alpha_{j,t} R_{j,t+1}, \sum_{j=1}^{n} \alpha_{j,t} = 1$。所有投资者的实际财富 W_t 都包括股票、债券、房地产等等。不幸的是，一阶条件包含无法观察的 $E_t U_{t+1}$，如果消费增长独立同分布，那么一阶条件可以写成

$$\delta E\left\{\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\ (R_{i,t+1} - R_{f,t})\right\} = 0 \tag{22a}$$

$$\delta\ [E\ (C_{t+1}/C_t)^{1-\gamma}]^{(\gamma-1/\psi)/(1-\gamma)} E\ \{\ (C_{t+1}/C_t)^{-\gamma} R_{f,t}^{*}\}\ = 1 \tag{22b}$$

这就避免了使用无法观察的市场收益，市场收益包括人力资本、家庭财富和股票市场财富等等（Kocherlakota（1996））。为了解释股权溢价，方程（22a）与标准（具有幂效用函数）的一阶条件相同。所需要的风险规避系数仍然很高。但是，如果给定 γ，我们就能在（22b）中选择 ψ 来与样本矩精确匹配。因此，在这些 *GEU* 偏好下不存在无风险利率之谜。

我们能对递归效用函数谈论得更为深入一些，这将为下一章要见到的跨期模型提供一些直觉。如果我们愿意接受如下假设：只有市场收益进入预算约束方程 $W_{t+1} = (1+R_{m,t+1})\ (W_t - C_t)$，那么 Campell、Lo 和 Mackinlay（1997）指出消费和资产收益服从对数正态分布情形的递归效用函数给出了如下结果：

$$\ln M_{t+1} = \frac{(1-\gamma)}{(1-\psi)} \Delta c_{t+1} - \frac{(1-\gamma\psi)}{(1-\psi)} r_{m,t+1}^{*} \tag{22c}$$

$$r_{f,t}^{*} = -\ln\delta + \frac{(\lambda-1)}{2}\sigma_m^2 - \frac{\lambda}{2\psi^2}\sigma^2\ (\Delta c)\ + \frac{1}{\psi} E_t \Delta c_{t+1} \tag{22d}$$

$$E_t\ (r_{i,.t+1}^{*} - r_{f,t}^{*})\ + \frac{\sigma_t^2\ (r_t^{*})}{2} = \frac{\lambda}{\psi}\sigma_t\ (r_i^{*},\ \Delta c)\ +\ (1-\lambda)\ \sigma_i\ (r_i^{*},\ r_m^{*}) \tag{22d}$$

其中，$\lambda = (1-\gamma) / (1-1/\psi)$，小写字母表示取对数后的值（即 $\ln R^{*} \equiv r^{*}$）。这些方程表明递归效用函数模型包括具有幂效用函数的 CCAPM 模型（即 $\gamma = 1/\psi$ 和 $\lambda = 1$ 的情形）和标准 CAPM 模型（效用函数为幂效用函数，因为 $\gamma = 1$ 所以消费与财富成比例。$\gamma = 0$，上述方程中只有一项 $\lambda\sigma_t\ (r_i^{*},\ \Delta c)\ = \lambda\sigma_i\ (r_i^{*},\ r_m^{*})$，$r_m^{*}$ 是财富组合的收益）。然而，就一般情形而言，任意风险资产 i 的平均超额收益取决于消费增长和资产收益之间的协方差（CCAPM 模型）及市场收益与资产收益之间的协方差。

时变风险溢价

对欧拉方程（22a）和（22b）的估计给出了对这些 SDF 模型进行检验的一种方法，但是，它不能给出风险溢价的时间序列——因为这一点，我们需要可观察变量的简化形式解和以协方差为条件的时变风险模型。我们来看看如何用 GARCH 模型来模拟时变风险溢价。SDF 模型可以表示为实际值的形式（方程 22d 和方程 22e）或名义值的形式。名义预算约束为

$$W_{t+1}^{nom} = (1+R_{m,t+1}^{nom})\ (W_{t+1}^{nom} - P_t C_t)$$

或

$$W_{t+1} = \frac{(1+R_{m,t+1}^{nom})}{(1+\pi_{t+1})}\ (W_t - C_t)$$

其中，P = 消费价格指数，W^{nom} 为名义财富，W 为实际财富，$(1+R_{t+1}) = (1+R_{t+1}^{nom}) / (1+\pi_{t+1})$，$(1+\pi_{t+1}) = P_{t+1}/P_t$。当使用名义收益时，Epstein - Zin 效用函数和联合正态分布（Campell、Lo 和 MacKinlay（1997），Smith、Sorensen 和 Wickens（2003））给出了

$$E_t\ (R_{i,t+1}^{nom}-R_{f,t}^{nom})\ +\frac{1}{2}\sigma_t^2\ (R_i^{nom})\ =-\frac{(1-\gamma)}{(1-\psi)}\mathrm{Cov}_t\ (\Delta c_{t+1},\ R_{i,t+1}^{nom})$$
$$-\frac{\psi\ (1-\gamma)}{(1-\psi)}\mathrm{Cov}_t\ (\pi_{t+1},\ R_{i,t+1}^{nom})$$
$$+\frac{(1-\psi\gamma)}{(1-\psi)}\mathrm{Cov}_t\ (R_{m,t+1}^{nom},\ R_{i,t+1}^{nom})$$

试图使用上式的一个问题在于，如何找到一个合适的指标来对包含所有资产的市场组合收益进行度量。如果我们（非常英雄式地）假设市场投资组合收益包括固定权重的平均股本收益和无风险利率，那么

$$R_{m,t+1}^{nom}=\theta R_{i,t+1}^{nom}+\ (1-\theta)\ R_{f,t}^{nom}$$

且

$$E_t\ (R_{i,t+1}^{nom}-R_{f,t}^{nom})\ +\left[\frac{1}{2}-\frac{\theta\ (1-\psi\gamma)}{(1-\psi)}\right]\sigma_t^2\ (R_{i,t+1}^{nom})\ =-\frac{(1-\gamma)}{(1-\psi)}\mathrm{Cov}_t\ (\Delta c_{t+1},\ R_{i,t+1}^{nom})$$
$$-\frac{\psi\ (1-\gamma)}{(1-\psi)}\mathrm{Cov}_t\ (\pi_{t+1},\ R_{i,t+1}^{nom})$$

注意，$c_t\equiv \ln C_t$ 为实际消费，但所有变量都是名义值。令$\beta_2=-(1-\gamma)/(1-\psi)$ 和 $\beta_3=-\psi(1-\gamma)\ /\ (1-\psi)$，那么对上述方程进行估计就给出了 $\gamma=\beta_2-\beta_3+1$ 和 $\psi=\beta_3/\beta_2$ 的估计值，$\sigma_t\ (R_{i,t+1}^{nom})$ 的系数没有任何限制。

因子模型

因子模型假设 SDF 线性地依赖于 k 个因子 $f_{i,t}$。

$$\ln M_{t+1}=\sum_{j=1}^{k}\beta_j f_{j,t+1}$$

如果因子和股本收益（记为 i）服从联合对数正态分布，那么 SDF 方法给出了无套利条件

$$E_t\ (r_{i,t+1}^*-r_{f,t}^*)\ +\frac{\sigma_t^2\ (r_i^*)}{2}=\sum_{j=1}^{k}\beta_j\sigma_t\ (r_i^*,\ f_j) \tag{22f}$$

这是最为一般的 SDF 模型，因为 b_j 没有任何限制，但实践中会遇到因子选择问题。具有幂效用函数或 Esptein - Zin 效用函数（递归 CES 效用函数）的 CCAPM 模型给出了方差项的特殊系数（如与方程（22e）一样）。“静态” CAPM 模型可以解释为试图解释资产 i 预期超额收益的限制性很强的 SDF 模型，在这个模型中我们只有市场收益一个因子，协方差项为 $\gamma\sigma_i\ (r_i^*,\ r_m)$。如果我们试图解释一些加总股票指数的超额收益 $E_t\ (r_{i,t+1}^*-r_{f,t}^*)$，我们也假设市场投资组合收益由固定权重的平均股本收益和无风险利率组成，那么协方差 $\gamma\sigma_t\ (r_i^*,\ r_m^*)$ 项变为方差项 $\gamma\sigma_t\ (r_m^*,\ r_m^*)$。这些限制性假设允许我们将第 5 章中的静态 CAPM 模型（5.4 节）解释为 SDF 模型。

一般而言，SDF 模型的关键结果是预期超额收益依赖于一个或几个协方差项（它们可能是时变的）。这对 CCAPM 模型和静态 CAPM 模型而言也是成立的。因为上世纪 90 年出现了 ARCH 模型和 GARCH 模型（见第 29 章），所以涌现了大量考察资产收益（特别是对一些加总股票指数的收益）和方差的研究，但是，一般而言这些结果与 SDF 模型的无套利条件不一致，这意味着我们使用了协方差项。从实证的角度来看，使用条件方差来解释预期收益是非常普遍的做法，因为它只需要使用单变量 ARCH/GARCH 过程，而应用 SDF 方法则要求使用多变量过程。当然，如果所确定的 SDF 模型经验效果很糟糕，那么作为数据的一种描述方式，方差项的使用将是不合适的。

欧拉方程方法不会给出风险溢价的时间序列，因为风险溢价在欧拉方程的一阶条件中是隐含变量，而在上述 SDF 模型中风险溢价是显示变量。在欧拉方程和简化形式方法中，我们已经作了一些辅助简化假设（如幂效用函数）。当需要估计的参数很多时，尽管存在“收敛”问题，SDF 模型中风险溢价的协方差序列还是可以用多变量的 ARCH/GARCH 模型来获得。根

据 SDF 模型，由 GARCH 模型得到的条件方差在预期收益的均值方程中（例如，参见（22e））应是统计显著的，因此，这些模型又被称为多变量 GARCH－M 模型，即 MGM 模型。

Smith、Sorensen 和 Wickens（2003）使用英国和美国 1975（6）～2001（12）期间的月度数据，与多变量 GARCH 过程合在一起，来估计与总股本收益 E_t（$R^{nom}_{i,t+1}-R^{nom}_{f,t}$）相对应的均值收益方程，GARCH 过程使用的向量为 $x=\{R^{nom}_{i,t+1}-R^{nom}_{f,t},\ \pi_{t+1},\ \Delta c_{t+1},\ \Delta y_{t+1}\}'$，其中 y_t（总产出的对数）是作为一个额外的“因子”包括进来的。与预期收益方程相对应的 MGM 模型可以写成

$$x_{t+1}=\alpha+\Gamma x_t+Bg_t+\varepsilon_{t+1}$$

其中，$\varepsilon_{t+1}\mid\Omega_t\sim D(0,\ H)$，$g_t=vec(H_{t+1})$。*vec* 算子将一个对称矩阵的下三角部分转换成一个向量，“*D*”代表“分布”，通常使用的分布要么是正态分布，要么是学生分布（嵌入肥尾）(参见第 29 章)。平均收益方程也包含一个反映 1987 年 10 月股灾的虚拟变量。在一般的 SDF 模型中，Γ 的第一行应该为零，B 的第一行为（$-1/2,\ \beta_{11},\ \beta_{12},\ \beta_{13},\ \cdots$），上述方程代表的是（22f）。

协方差项 H_{t+1} 可以用多变量 GARCH 模型来估计（参见第 29 章），Smith、Sorensen 和 Wickens（2003）使用的是带限制的 VECM（1，1）－BEKK 模型（Engle 和 Kroner 1995）：

$$H_{t+1}=V'V+A'(H_t-V'V)A+F'(\varepsilon_t\varepsilon_t'-V'V)F$$

在一般的 SDF 模型中，协方差项是股票市场收益 $R^{nom}_{i,t+1}$ 与变量 $\{\pi_{t+1},\ \Delta c_{t+1},\ \Delta y_{t+1}\}$ 之间的协方差。用 GARCH 模型协方差项进行建模的时变风险溢价可以解释超额股票收益运动的一个部分，但在 Esptein－Zin CCAPM 模型中 $\gamma=\beta_2-\beta_3+1$ 和 $\psi=\beta_2/\beta_1$ 的估计值都是不可接受的。例如，就美国而言，$\hat{\gamma}=783$，$\hat{\psi}=-118$；对英国而言，$\hat{\gamma}=844$，$\hat{\psi}=-4.5$。幂效用函数模型给出的结果与 *Esptein－Zin* 模型类似，因此数据不支持 CCAPM 模型的两种变体。在一般的 SDF 模型中（在平均收益方程中，协方差项的大小和符号都没有任何限制），Smith、Sorensen 和 Wickens（2003）发现，在平均收益方程中与消费有关的时变协方差项（美国 t 统计量≈2.9，英国 t 统计量≈1.57）和与通货膨胀有关的时变协方差项（美国 t 统计量≈1.7，英国 t 统计量≈2.8）恰好在统计上是显著的，但与产出有关的时变协方差和市场收益的方差却不显著（市场收益的方差排除了静态 CAPM 模型的可能性）。由 SDF 风险溢价所解释的超额收益运动的比例约为 4%。

整体而言，用可观察因子（如通货膨胀、消费和产出）对时变协方差进行建模和使用英国或美国月度数据来估计（带限制的）GARCH 模型都是非常困难的，即使用很流行的“以数据为基础的”SDF 模型也很难发现预期股票收益的协方差效应会在统计上显著。

不使用消费数据的 SDF 模型

Campell（1993）说明了如何不用消费数据而只用收益数据来表示上述超额收益方程。思想是基于财富约束的对数线性化公式

$$\Delta\omega_{t+1}\approx r^{*}_{m,t+1}+k+\left(1-\frac{1}{\rho}\right)(c_t-\omega_t)$$

其中，$r^{*}_{m,t+1}=\ln R^{*}_{m,t+1}$，其余类似，$\rho=1-\exp(\overline{c-\omega})$，$k$ 是线性化后的常数。（注意，我们在这里定义的 ρ 与 Campell 的记号一致。）将上述方程向前迭代并使用 $r^{*}_{f,t}$ 和 $r^{*}_{i,t+1}$ 的上述对数线性方程，Campell 证明了

$$c_t-\omega_t=(1-\psi)\ E_t\Big[\sum_{j=1}^{\infty}\rho^j r^{*}_{m,t+1}\Big]+k^{*}$$

k^{*} 在这里取决于 $\mathrm{Cov}_t(r^{*}_m,\ \Delta c)$。很明显，$\psi>1$ 意味着预期收益的替代效应占优于收入效应，消费财富比随着 $Er^{*}_{m,t+j}$ 的增加而递减（当 $\psi<1$ 时，结论与之相反）。如果 $\psi=1$，那么消费财富比与预期收益独立。如果再向前推进一点，我们就能获得预期收益方程的一种“漂亮”直觉。由上述方程可以直接导出

$$c_{t+1}-E_tc_{t+1}=(w_{t+1}-E_tw_{t+1})+(1-\psi)(E_{t+1}-E_t)\left[\sum_{j=1}^{\infty}\rho^jr^*_{m,t+j+1}\right]$$
$$=(r^*_{m,t+1}-E_tr^*_{m,t+1})+(1-\psi)(E_{t+1}-E_t)\left[\sum_{j=1}^{\infty}\rho^jr^*_{m,t+j+1}\right]$$

最后一步替换来自于对数线性预算约束。我们将再一次看到 $\psi>1$ 或 $\psi<1$ 在决定未来预期收益的修正（$E_{t+1}-E_t$）对消费变化的影响方面的关键性作用。然而，更为重要的是上述方程意味着任意资产 i 与消费之间的协方差可以表示为收益 i 与市场收益（即财富组合）之间的协方差，该协方差可以表示为

$$E_t(r^*_{i,t+1}-r^*_{f,t})+\frac{\sigma_t^2(r_i^*)}{2}=\gamma\sigma_t(r_i^*,r_m^*)+(\gamma-1)\sigma_t(r_i^*,h_m)$$

其中

$$\sigma(r_i^*,h_m)\equiv\mathrm{Cov}_t\left(r^*_{i,t+1},E_{t+1}\left[\sum_{j=1}^{\infty}\rho^jr^*_{m,t+1+j}\right]-E_t\left[\sum_{j=1}^{\infty}\rho^jr_{m,t+1+j}\right]\right)$$

我们已经将 r_i^* 与消费增长之间的协方差替换为（i）r_i^* 与市场收益之间的协方差和（*ii*）r_i^* 与未来市场收益消息之间的协方差，而从在超额收益方程中消失了。

上述方程是 Merton（1973）连续时间模型的一个离散版本，模型中的预期资产收益依赖于能代表投资机会集合变化的“套期保值组合”（即协方差项 $\sigma_t(r_i^*,h_m)$）。另外，由上述方程明显可以看出在哪些特殊条件下资产收益与市场组合收益之间的协方差能成为资产定价的一个充分统计量（即模型简化为静态 CAPM 模型）。这些条件是：

（ⅰ）投资机会集合为常数，$\sigma_t(r_i^*,h_m)=0$；

（ⅱ）相对风险规避系数 $\gamma=1$；

（ⅲ）r_m 服从单变量随机过程，于是未来收益方面的消息与当前收益完全相关（因此，$\sigma_t(r_i^*,r_m)$包含与 $\sigma_t(r_i^*,h_m)$ 有关的所有信息）。

正如我们在下一章将要看到的那样，投资机会集合为常数（收益独立同分布）或 $\gamma=1$（对数效用函数）的假设是跨期模型非常重要的“特殊情形”。

对市场收益而言上述方程变成

$$E_t(r^*_{m,t+1}-r^*_{f,t})+\frac{\sigma_t^2(r_m^*)}{2}=\gamma\sigma_t^2(r_m^*)+(\gamma-1)\sigma_t(r_m^*,h_m)$$

注意，如果（ⅰ）$\gamma=1$ 或（ⅱ）$\gamma\neq1$ 且市场收益不可预测（即 $\sigma_t(r_m^*,h_m)=0$），那么可以用历史夏普比来对风险溢价进行度量。如果 $\gamma>1$ 且 $\sigma_t(r_m^*,h_m)<0$（即长期均值反转），那么夏普比夸大了风险溢价。

国际 CAPM 模型

如果将 Campell 模型用于对外投资，那么用本国货币表示的实际收益取决于国内价格水平和实际汇率的变化。如果收益是名义收益，那么（本币）通货膨胀是需要考虑的另外一个变量。如果我们以国家 1 的货币为记价货币，那么 j 国投资者的实际收益为（见第 1 章）

$$R^*_{p,t+1}=R^{nom}_{p,t+1}\frac{P_t^1}{P_{t+1}^1}\frac{Q_t}{Q_{t+1}}$$

其中，$R^{nom,1}_{p,t+1}=j$ 国投资者的投资组合用国家 1 货币表示的名义收益，$P_t^1=P_t^1$ 用国家 1 货币表示的价格水平，$Q_t=$ 实际汇率。取对数后上式变成

$$r^*_{p,t+1}=r^{nom,1}_{p,t+1}-\pi^1_{t+1}-\Delta q_{t+1}$$

其中 $\pi\equiv\ln(P_{t+1}^1/P_t^1)$。如果购买力平价理论（PPP）成立，那么 j 国投资者的实际收益就可以简单地表示为，用货币 1 表示的收益减去用货币 1 表示的通货膨胀率。但是，如果 PPP 理论不成立，那么实际收益就取决于实际汇率。国际投资者的（实际）预算约束为

$$W_{t+1}=R^*_{p,t+1}(W_t-C_t)$$

可以用上面的方法对此进行线性化处理（Campell（1993）），然后对预期收益方程中的消费进

行替换。在递归效用函数的假设下，资产的预期收益方程为（Ng（2004））

$$E_t\ (r_{i,t+1}^{nom}-r_{f,t}^{nom})\ +\frac{\sigma_t^2\ (r_i^{nom})}{2}=\gamma V_{i,m}+\ (1-\gamma)\ V_{i,q}+V_{i,\pi}+\ (\gamma-1)\ (V_{i,m}-V_{i,hq})$$

其中 $V_{i,m}=\mathrm{Cov}_t\ (r_{i,t+1}^{nom,1},\ r_{m,t+1}^{nom,1}-\pi_{t+1}^1)$ 是资产 i 收益与实际市场收益之间的协方差。另外，$V_{i,q}=\mathrm{Cov}_t\ (r_{i,t+1}^{nom,1},\ \Delta q_{t+1})$，$V_{i,hm}=\mathrm{Cov}\ (r_{i,t+1}^{nom,1},\ h_m)$，$h_m$ 是未来实际市场收益方面的消息，即

$$h_m=\ (E_{t+1}-E_t)\ \sum{}_{k=1}^{\infty}\rho^k\ (r_{m,t+k+1}^{nom,1}-\pi_{t+k+1}^1)$$

$V_{i,hq}=\mathrm{Cov}\ (r_{i,t+1}^{nom,1},\ h_q)$，$h_q$ 的定义与 h_m 相似，但它是与未来价值 Δq_{t+k+1} 有关的“消息”。

由上述方程可以明显看出，“国际 CAPM 模型”在结构上与 Campell（1993）的“国内 CAPM 模型”相似，但是，由于国际变量（Δq_{t+1}）与通货膨胀率 π_{t+1} 方面的原因，国际 CAPM 模型中多了一个协方差项，因为收益是名义收益。“V”项上的系数限制也是很明显的。$V_{i,hm}$ 和 $V_{i,hq}$ 为跨期对冲项，但是，如果“国内”股票收益和实际汇率都不可预测，那么这些变量都应等于零，模型变为“静态”ICAPM 模型（模型中的收益只取决于资产收益与市场收益、实际汇率和通货膨胀率之间的协方差）。当然，当 PPP 理论成立时，$V_{i,q}=V_{i,hq}=0$，模型简化为 Campell（1993）的“国内 CAPM 模型”。最后，如果 $\gamma=1$，那么标准（静态）CAPM 模型是适用的。Ng（2004）针对美国、英国、日本和德国（1978 年 7 月 ~ 1998 年 4 月）的国内股票收益和外汇收益数据来运用国际 CAPM 模型（没有消费）。市场收益为世界股票指数的收益。h_m 和 h_q 中“突发事件可以通过对包含这些（和其他）变量的 Var 模型进行估计”来进行测度，从而 z_t 和未来突发事件依赖于 Var 模型中待估的协方差矩阵（参见第 12 章）。我们也假设预期股票收益和外汇收益 $r_{i,t+1}^{nom}$ 依赖于变量 z。

Ng（2004）发现，跨期对冲项 $V_{i,hm}$ 和 $V_{i,hq}$ 在确定横截面资产收益方面（在统计上）都很重要。然而，他也发现 $\gamma\ (V_{i,hm}-V_{i,hq})$ 项与 $\gamma V_{i,m}+\ (1-\gamma)\ V_{i,q}+V_{i,\pi}$ 项成比例，于是后者足以解释横截面收益，模型简化为“静态”国际 CAPM 模型。因此，从实证角度来看动态对冲项的作用是有限的，模型也不能解释 Fama 和 French（1998）中高账面市值比国家投资组合的平均收益。

14.5　习惯形成

现在，我们转为讨论另外一种形式的不可分离性——习惯持续性。在股权溢价之谜和无风险利率之谜的分析中，习惯持续性不可分离性质非常有用。在习惯形成模型中，并不是消费的绝对水平很重要，而是相对于历史基准水平的消费很重要。这似乎是一个合理的起点，因为“因果内省”建议人们担心那些相对于历史消费水平的消费减少，即使消费有所下降，消费的绝对水平在过去 20 年中也可能是处于第二或第三的位置。另外，社会保险合同（如来自于汽车、个人意外事件和失业保险等的偿付）很少对全部价值提供保险，因此它们只能预防价值的下降。

习惯持续模型（habit - persistence model）通常都假设代表性投资者存在，模型中效用函数的形式具有不可分离的形式，效用函数为 $U\ (C_t,\ Z_t)$，$Z_t=f\ (C_{t-1},\ C_{t-2},\ \cdots)$。$U\ (C_t,\ Z_t)$ 的准确形式可能有所变化，如下述效用函数。个体习惯持续性可以表示为

$$U=\frac{(C_{i,t}-\lambda C_{i,t-1})^{1-\gamma}}{1-\gamma}\quad \lambda>0 \tag{23}$$

“和你的富裕邻居保持同样生活方式或质量（keeping up with the Jones's）”可以表示为

$$U=\frac{C_{i,t}^{1-\gamma}C_t^{\lambda}C_{t-1}^{\lambda}}{1-\gamma} \tag{24}$$

其中，$C_{i,t}$ 是个体消费，C_t 是总消费（Constantinides（1990），Heaton（1995），Abel（1995），Gali（1994））。在使用这些模型时存在一个问题：一阶条件包含 θE ｛（U'_{t+1}/U'_t）（$R_{i,t+1}$ −

$R_{f,t})\}=0$。但 U'_{t+1} 依赖于投资者预测未来消费增长的能力，除非我们做出消费增长独立同分布的辅助假设，否则 U'_{t+1} 无法观测。所有这些模型都不能同时解释股权溢价之谜和无风险利率之谜。

我们将更为深入地讨论 Campell 和 Cochrane（1999）的习惯持续模型，期望获得一些直觉：为什么在标准偏好和低“幂系数”γ 情形下它能解释股权溢价之谜，也能对低水平值且波动性小的利率进行预测。原因在于，局部凹性 $CU''(C)/U'(C)$ 不仅依赖于 γ 还依赖于消费高于（或低于）习惯消费水平的程度。所以，局部凹性或“风险规避系数”在衰退中也可能很高（即当前消费处于习惯消费水平之下）。低的幂系数给出了低水平值但波动性小的利率（即低的跨期替代弹性）。但是，投资者仍然会发现股票是有风险的，并不是因为他们的消费很低而是因为他们不喜欢衰退，这就给出了高的局部凹性和高的风险规避系数。换一种方式说，投资者不喜欢股票是因为他们在衰退中得到的偿付很低。

Campell - Cochrane：习惯持续 CCAPM 模型

正如上面所注意到的那样，Campell - Cochrane 模型的一个关键因素是幂效用函数的（局部）凹性依赖于相对于近期习惯性消费与当前消费的相对水平。决定效用函数的不是高或低的绝对消费水平，而是相对于近期平均值而言高或低的当前消费水平。因此，“新的”状态变量是“历史消费或习惯 X_t 的加权平均”，幂效用函数关于 C_t 和 X_t 不具有可分性（因此，效用函数不能写成 $U_t(C_t, X_t)=U(C_t)+U(X_t)$ 的形式）：

$$U_t=\sum_{t=0}^{\infty}\theta^t\frac{(C_t-X_t)^{1-\gamma}-1}{1-\gamma}\tag{25}$$

其中，$U_t(C_t, X_t)=[(C_t-X_t)^{1-\gamma}-1]/(1-\gamma)$。效用函数体现了“和你的富裕邻居保持同样生活方式或质量”的思想，因为个体习惯都由个体的历史消费来决定（Abel（1990））。此外，

$$U_c(C_t, X_t)=(C_t-X_t)^{-\gamma}=(S_tC_t)^{-\gamma}\tag{26}$$

其中，$S_t\equiv(C_t-X_t)/C_t=1-X_t/C_t$，$S_t$ 为“剩余消费比率”（surplus consumption ratio），S_t 是衰退指标：当 C_t 朝以前的平均水平下降时，S_t 会上升。效用函数的凹性（松散地说就是风险规避系数）为

$$\eta_t=-\frac{C_tU_{cc}(C_t, X_t)}{U_c(C_t, X_t)}=\frac{\gamma}{S_t}\tag{27}$$

因此，即使“幂系数”γ 很小（如），当当前消费接近于以前的“习惯水平”X_t 时，η_t 也会很大。还不仅仅如此，随着衰退时期的来临，S_t 下降，η_t 上升，这意味着衰退时期中的风险规避系数会增加。模型的夏普比为（见方程（4））

$$\frac{E_tR_{t+1}-r}{\sigma_t(R_{t+1})}=\eta_t\sigma_t(\Delta c_{t+1})\rho_t(\Delta c_{t+1}, R_{t+1})\tag{28}$$

其中，$r=\ln R_f^*$。高平均 η_t 意味着我们能解释高平均夏普比。此外，即使 σ_t 和 ρ_t 为常数，因为 η_t 在衰退期中上升，模型也能解释夏普比的逆周期模式。SDF 为

$$M_{t+1}\equiv\theta\frac{U_c(C_{t+1}, X_{t+1})}{U_c(C_t, X_t)}=\theta\left(\frac{C_{t+1}-X_{t+1}}{C_t-X_t}\right)^{-\gamma}=\theta\left(\frac{S_{t+1}C_{t+1}}{S_tC_t}\right)^{-\gamma}\tag{29}$$

当 γ“很小”且等于 2 时，我们对跨期替代的“厌恶”程度很低，因此，当利率的变化很小时，消费者愿意在 t 和 $t+1$ 之间转移消费（而“大的”η_t“给出了高的股权溢价”）。事实上在这个模型中，

$$r_t=-\ln E_t(M_{t+1})=-\ln\theta+\gamma g-\frac{1}{2}\left(\frac{\gamma}{S}\right)^2\sigma_t^2(\Delta c)\tag{30a}$$

$$\bar{S}=\sigma(\Delta c)\sqrt{\frac{\gamma}{1-\phi}}$$

$$g = E_t\ (\Delta c_{t+1}) \tag{30b}$$

$s = \ln S$，参数 ϕ（$=0.87$）表示持续程度。由（30a）可以直接看出“小的”$\gamma = 2$ 意味着对消费增长 g 的变化有细微反应。

在保持利率相对稳定的过程中存在两种相反的力量。当我们进入衰退期时，C_t 朝 X_t 的方向递减，边际效用 $U_c = (C-X)_t^{-\gamma}$ 递增。在前面的模型中，这导致了借贷行为（因此，当 U_c 很高时，你能消费得更多），驱动利率上升。模型中的相反力量是：在衰退期中风险规避系数 η_t 很高，这将导致谨慎储蓄的增加，而谨慎储蓄的增加倾向于降低利率。两种效应倾向于相互抵消，从而使利率接近于一个常数。

从本质来看，Campell – Cochrane 的聪明之处在于他们在 M_{t+1} 对消费增长的反应中引入了非线性关系（见（29））。剩余消费比的对数 s_{t+1} 服从缓慢的均值反转 AR（1）过程，对消费增长冲击 v_{t+1} 的非线性反应为 $\lambda\ (s_t)$：

$$\Delta c_{t+1} = g + v_{t+1} \quad v_{t+1} \sim niid\ (0,\ \sigma^2) \tag{31}$$

$$s_{t+1} = (1-\phi)\ \bar{s} + \phi_t s_t + \lambda\ (s_t)\ v_{t+1} \tag{32a}$$

其中

$$\lambda\ (s_t) = \frac{1}{\bar{S}}\sqrt{1-2\ (s_t-\bar{s})} - 1 \tag{32b}$$

（32）中相当复杂的过程保证了消费始终在习惯消费水平之上，因为 $S = e^s > 0$。将（29）中 M_{t+1} 取对数后的 $\ln M_{t+1}$ 代入（32），s_{t+1} 服从如下过程：

$$m_{t+1} = a + b\ (s_t)\ + d\ (s_t)\ \Delta c_{t+1} \tag{33}$$

$b\ (s_t)$ 和 $d\ (s_t)$ 是 s_t 和参数 θ，γ，ϕ 的非线性函数。因此，不仅 m_{t+1} 对消费增长 Δc_{t+1} 会产生反应（正如非习惯形成 CCAPM 模型那样），m_{t+1} 对 Δc_{t+1} 的敏感性还依赖于经济状态 s_t。

模型中相对收益（ER_i/ER_j）的情况怎样呢？当 $\gamma = 2$ 时，根据（29）中 M_{t+1} 的定义，ER_i/ER_j 较大的值来自于 $(S_{t+1}/S_t)^{-\gamma}$ 项，即 R_i 或 R_j 与相对于习惯的消费水平（“衰退”变量的一种）之间的协方差。在非习惯形成 CCAPM 模型中，相对收益由资产收益与消费增长本身之间的协方差而不是与基于近期平均水平的相对消费之间的协方差来决定。

校准与模拟

因为这个模型是校准模型，所以没有正规的统计检验，但可以用模拟方法来对模型进行评价，然后查看模型输出是否与真实世界数据所反映的典型事实相吻合。利用历史数据可以将潜在的“基础变量值”取为 $g = 1.89$，$\sigma = 1.9$，$r = 0.94$。参数 ϕ 选择为 0.87，因此，模型产生了价格红利比的自相关性，根据数据也能发现这种自相关性。此外，取 $\gamma = 2$，从而 $\bar{S} = 0.57$。应用 $v_{t+1} \sim niid\ (0,\ \sigma^2)$ 和各种已经计算出来的统计量对模型进行 10 000 次模拟。θ 的合理值取为 0.89，它能产生 0.94% 的（平均）利率和 0.5 的夏普比。根据模型模拟的结果，我们能获得股票价格 P_t 及收益的时间序列数据。将不同时期收益对价格红利比进行回归的结果表明，负的回归系数和 R^2 都是时期长度的增函数，这与数据中所发现的结果非常接近。因此，模型“模拟”了资产收益的可预测性、高夏普比、值小且波动性小的实际利率，这些都是实际数据所反映的典型事实。但要注意的是，要解释股权溢价就必须假设风险规避系数很大（也就是说，当 C_t 与习惯消费 X_t 很接近时，η_t 会非常的高）。然而，无风险利率与 η_t 的变化不一致，无风险利率依赖于 γ，而 γ 可以很小，于是它导致了变化很小的无风险利率。

消费的边际效用为：

$$U_c = C_t^{-\gamma}\left(\frac{C_t - X_t}{C_t}\right)^{-\gamma}$$

因此消费者不喜欢低消费（与标准模型中的情形一样），但他们也不喜欢衰退（即相对于近期消费水平而言消费很低的时候）。股票存在风险，部分原因是消费具有波动性（但波动性不是很大），主要原因在于，衰退时期中糟糕的股票收益倾向于发生（即股票收益与习惯变量

相关）。

Campell 和 Cochrane（1999）用上述模型阐述了为什么 CCAPM 模型相对于标准 CAPM 模型或条件 CAPM 模型的效果会更糟糕一些——那也是诸如红利价格比等预测变量也会对资产收益产生影响的原因（Cochrane（1996））。他们根据习惯形成模型生成了收益和消费的高频数据，因此模型中只有一个冲击。根据构造，消费增长、收益和 SDF 相互之间都是条件相关的且相关系数为正。这也是在很短时期内模型中非线性几乎消失的原因。然而，随着时间间隔的增加，消费增长、收益和折现因子对“冲击”的反应不同，因此它们之间就不会出现完全无条件相关的情形，例如价格红利比是未来收益的一种预测。标准 CAPM 模型和条件 CAPM 模型对横截面收益的预测效果比习惯形成 CCAPM 模型要好，即使习惯形成模型是正确的模型。这是市场收益与正确 SDF 之间条件相关性更好的一个原因，在当价格红利比发生改变后这种相关性会受到影响，而消费却不会受到影响。当然，这建立了消费习惯形成模型的表面情形，但并没有证明它是资产收益的正确模型。

14.6 股权溢价：更进一步的解释

不完备市场

不完备市场的一般均衡模型相当复杂，经常用数字方法来进行求解，模型的直觉也很难获得。尽管如此，我们在这一节中还是要对完备市场 SDF 模型与两期和无限期界的不完备市场模型进行比较。后者关注的中心是个体消费，而 SDF 模型通常要借助于“代表性投资者”假设，因此模型中的消费是总量消费，它与资产收益之间的相关性是预期收益的关键驱动因素。

假设个体能为自身提供保险，因此个体也能抵御异质性冲击。例如，面对失业，个体要么持有保险合同以期在“不利”状态下获得支付，要么拥有足够多的金融资源（储蓄）来“克服暴风雪天气”。因此，个体消费与总的人均消费类似。但是，我们在标准偏好下再一次遭遇了股权溢价之谜。

不完备市场文献保留了标准偏好，也注意到了如下事实：当市场不完备时，个体消费可能比人均消费有更大的变化。因此，个体消费与股票收益之间的相关性可能更高，从而低的风险规避系数可以解释所观察到的股权溢价。不完备市场上个体消费的变化也意味着个体为了预防消费冲击将进行更多的储蓄来为自己提供保险，储蓄的增加降低了平均无风险利率。因此，这些模型有可能同时解释股权溢价之谜的两个部分（Weil（1992））。

不完备市场的两期一般均衡模型还不足以捕获这种“自我保险”，但在一个很长的时期中“自我保险”是完全可能的。例如，如果我们只活两期（且没有任何遗产），那么第二年中负的收入冲击可以用更低的个体消费来完全反映，个体消费的变化程度有所增加。但是，如果我们期望活很多年，那么可以通过如下方式来匀滑消费：在“冲击时期”中进行少量储蓄，在收入高时进行更多储蓄。这种动态“自我保险”减弱了异质型风险对个体消费的影响。

Kocherlakota（1996）提供了不完备市场无限期界模型中动态自我保险的一个“纯粹”直观例子。与两期不完备市场情形相比，无限期界模型中的储蓄更少，因此该模型在生成无风险利率方面的效果要差一些。

在这个“模型”中，寿命无限长的个体有无限多个，他们的收入是随机的且相互独立（因此人均收入的变化为零）。个体不能通过签订保险合同的方式来预防个体收入的变化。他们能相互发放贷款，但每个人向他人的借款不超过 B 美元。因为（持久）收入 y_{min}/r 永远不会低于 B 美元，所以我们假设开始时的 B 美元借款不受约束。如果交易到达了均衡，那么对每个放款人而言必然存在一个借款人，于是 $y_{min}>0$。

与两期模型相比，在没有任何保险的无限期界模型中消费下降的概率更小一些。尽管受信

额度 y_{min}/r “缓冲” 了消费的下降，但只有长期负收入冲击才能产生这一限制条件。因此，个体不会面临任何约束，即使在不完备市场中自我保险所需要的额外储蓄也比两期模型低。因此，不完备市场无限期界模型所给出的消费水平与完备市场模型中所发现的消费水平较为接近，所以不完备市场无限期界模型也不能 “拟合” 所观察到的 “值小” 的无风险利率。

在上述模型中，收入冲击是平稳的，最后会完全消失。在 Constandinides 和 Duffie（1996）的不完备市场模型中个体劳动收入冲击具有持久性，无限期界动态自我保险行为不可能发生，因为持久收入冲击在消费中已经完全反映出来了。如果持久收入降低 x 美元，那么为了匀滑消费，每一期的消费也必须下降 x 美元。因此，与完备市场情形相比，这个模型中的无风险利率会更低一些。

Constandinides 和 Duffie（1996）的模型非常巧妙，在产生任何股权溢价和低且相对稳定的无风险利率时都是 “方向反转的”（reverse engineered），模型要求无交易成本和借贷约束。正如我们已经看到的那样，与资产收益不相关的异质型风险对均衡收益没有任何影响。如果我们使异质型风险与市场收益相关，那么个体之间可以通过相互交易来 “消除” 这种相关性。例如，如果 A 在股市繁荣时获得了更多的劳动收入，当股市萧条时 B 获得了更多的劳动收入，那么 A 将资产卖给 B，这种交易将消除个体收入冲击与资产收益之间的相关性，从而对资产价格没有任何效应。

C－D 模型产生 “高风险”（风险规避系数低）的方式非常微妙，基本上取决于边际效用的非线性性。收入冲击与资产收益无关，因此它们能用相互交易的方式进行抵消。但是，非线性的边际效用将这些冲击转化为与资产收益相关的边际效用冲击，从而对均衡收益产生影响。使用 C－D 模型的结果是，当市场下跌时，异质型冲击的方差将上升。

通过如下方式可以看出一些基本直觉（Cochrane（2000））：对通常的一阶条件在 N 个个体上（$j=1, 2, \cdots, N$）进行加总，

$$0=E\ \{S_N\ [\ (C_{t+1}^{(j)}/C_t^{(j)})^{-\gamma}]\ (R_{t+1}-R_{ft})\} \tag{34}$$

其中，$S_N=\ (1/N)\sum_{j=1}^{N}[(C_{t+1}^{(j)}/C_t^{(j)})^{-\gamma}]$ 是对个体 j 进行求和的结果，$C^{(j)}$ 是个体 j 的消费。如果消费增长的横截面变化服从对数正态分布，那么

$$0=E\left\{\exp\left[-\gamma\ (S_N\Delta c_{t+1}^{(j)})\ +\frac{\gamma^2}{2}\sigma_N^2\ (\Delta c_{t+1}^{(j)})\right]\ (R_{t+1}-R_{ft})\right\} \tag{35}$$

其中，σ_N^2（·）表示消费增长在所有个体上的变化。如果 σ_N^2（·）与资产收益相关，那么该项引入了另外的 “价格” 风险。因为我们可以用 C－D 模型来拟合典型事实，所以模型的有效性与关键假设紧密相连：（i）个体劳动收入冲击是持久的；（ii）当市场收益更低时，个体收入的横截面不确定性是否更高；也就是说，当收入冲击也很 “糟糕” 时你是否要求一个高股权溢价，因为股票收益很低。Heaton 和 Lucas（1996）及 Deaton 和 Paxson（1994）中的证据表明，美国的横截面数据不支持 C－D 模型的假设（如无法分散的收入冲击的自相关性约为 0.5，而不是 1）。

C－D 模型和各种习惯持久型模型，尽管在结构上有很大的差异，但都有一个相同的部分。当消费者已经感觉很 “糟糕” 时，他们不会担心收益和真实财富的下降。在习惯持久型模型中，相对于近期消费而言消费低的日子就是坏日子，而在 C－D 模型中只有个体劳动市场风险高时才是 “坏” 日子。宽泛地讲，两种模型中所增加的状态变量可以描述为 “衰退变量”。

交易成本

一种可能的情形是股票交易相对于（无风险）债券交易而言的其他成本使得股票与债券之间的收益差距变大。假设债券交易没有任何成本。在无限长的时期中，红利为常数，股票收益为 $R_S=D/P$。可以证明，如果永远持有股票，那么股权溢价的上界（Kocherlakota（1996））为

$$R_S-R_f=\tau R_f$$

其中，τ 是股票交易的比例成本（这是相对于债券而言的成本，假设债券的交易成本为零）。因此，对 $R_f=1\%$ 而言，为了解释6%的股权溢价，增加的交易成本必须为600%。显然，如果个体有限制性借贷约束，那么在出现负的收入冲击后在不售卖股票的条件下（即你永远持有股票）你是无法匀滑消费的。如果这种事情频繁发生，那么股票交易就变得很昂贵，你将需要一个更高的股权溢价。

借贷约束有降低无风险利率的潜力（与无借贷约束的情形相比较）。如果人们的收入冲击为负且有限制性借贷约束，那么他们就不能借到任何东西。因此，假定债券（贷款）市场上需求必须等于供给，那么放款人比借款人要多，利率会更低一些（与无借款约束的模型相比）。但是，那些面临借款约束的人可以通过出售股票来获得资金并利用这笔资金来匀滑消费。然而，如果他们在股票市场上也受到约束（即没有股票），那么股票的供给就会超过需求（与无约束的情形相比），因此股票收益会下降（股权溢价在一些个体收到约束的模型中应相等，这与“无约束”的标准模型一样（Kocherlakota（1996）））。

14.7 小结

* CCAPM 模型与非常短时期（如至多 1 个月）内股票、债券和外汇收益的不可预测性是一致的。当消费的条件预测值和收益的波动率随着时间发生变化时（风险规避系数也可能发生变化），CCAPM 模型允许较长时期上的资产收益存在可预测性。然而，在实证研究中，收益可预测性回归方程中的变量通常不会与 CCAPM 模型中的矩直接相关。

* 具有幂效用函数（依赖于消费水平）的 CCAPM 模型不能解释所观察到的高股权溢价（即股权溢价之谜）和低均值低波动性的无风险利率（即无风险利率之谜）。

* 股权溢价之谜也说明 CCAPM 模型在解释所观察到的夏普比（比如说，标准普尔 500 指数）、几种资产的 H - J 边界、CCAPM 欧拉方程在横截面和时间序列上的行为方面都是失败的。

* 使用 SDF 模型来试图“解决”股权溢价之谜取得了一些进展，但仍不是很成功，最为著名方的法是在效用函数中引入其他状态变量（如财富、消费习惯等）。

* 使用不完备市场的方法来“解决”股权溢价之谜也是可能的，但所要求的个体收入持久冲击的假设不能为大家所广泛接受。

附录：Hansen - Jagannathan 边界

考虑 M_t 在 N 由种资产偿付所构成集合 X_t（$N\times 1$）上的回归投影：

$$M_t-E(M)=(X_t-E(X))'\beta+\varepsilon_t \tag{A1}$$

M_t 为一个指标，而 X 和 β 都是（$N\times 1$）向量。由（A1）和 $E(\varepsilon_t)=E(\varepsilon_t X)=0$ 的事实，

$$(M_t-E(M))(X_t-E(X))=(X_t-E(X))(X_t-E(X))'\beta+(X_t-E(X))\varepsilon_t \tag{A2}$$

因此

$$E(M_t X)=E(M)E(X)-\sum\beta \tag{A3}$$

其中，$\sum=\mathrm{Cov}(X,X')$。因为 M 能正确地对所有资产进行定价，所以 $P=E(M_t X)$，由（A3）可以解出 β：

$$\beta=\sum{}^{-1}[P-E(M)E(X)] \tag{A4}$$

根据（A1），如果再注意到如下事实：根据构造，回归方程中 ε_t 的 X_t 与无关，那么

$$\sigma^2(M) = \sigma^2\{(X_t - E(X))'\beta\} + \sigma^2(\varepsilon) = \sigma^2(z) + \sigma^2(\varepsilon) \tag{A5}$$

其中，$z = X_t - E(X)'\beta$。因此 $Ez = 0$ 且

$$\sigma^2(z) \equiv E(z'z) = \beta' \sum \beta \tag{A6}$$

将（A4）中的β代入，

$$\sigma^2(z) = [P - E(M)E(X)]' \sum{}^{-1} [P - E(M)E(X)] \tag{A7}$$

将（A7）代入（A4），注意$\sigma^2(\varepsilon) > 0$就得到正文中所发现的不等式：

$$\sigma^2(M) \geqslant [P - E(M)E(X)]' \sum{}^{-1} [P - E(M)E(X)] \tag{A8}$$

除以P之后（点态意义），向量P变为元素全是 1 的向量 e1，X变为总收益向量R^*，因此

$$\sigma^2(M) \geqslant [\mathrm{e1} - E(M)E(R^*)]' \sum{}_R^{-1} [\mathrm{e1} - E(M)E(R^*)] \tag{A9}$$

其中，$\sum_R$是收益的方差协方差矩阵。当只有一种风险资产时，我们令 e1 = 1，$\sum_R = \sigma^2(R^*)$。假设$E(M) = 1/R_f^*$，(A9) 就简化为正文中方程（7）Hansen – Jangannathan 边界的夏普比版本：

$$\sigma(M)/E(M) \geqslant E(R^* - R_f^*)/\sigma(R^*) \tag{A10}$$

第15章

跨期资产配置：理论

目的

* 求解两期消费－投资组合模型，模型中代表性投资者的效用函数具有时间可分离性且只依赖于消费。资产可以用来将财富从一个时期转移到下一时期（劳动收入不进入预算约束）。代表性投资者同时选择合意的消费水平和风险资产及无风险资产的持有数量。

* 说明在两期模型的普通解中消费和资产的份额取决于当前财富水平和下一期风险资产预期收益的预测值（以及无风险利率和风规避系数）。

* 将模型扩展到多期情形，代表性投资者在时刻决定最优消费路径（C_t，C_{t+1}，…，C_T）和未来各期中最优资产比例（α_t，α_{t+1}，…，α_T）。解的一般形式（具有时间可分离性）与两期情形相同，消费与资产份额取决于当前财富水平和未来预期收益。

* 对幂效用函数和独立同分布收益的特殊情形而言，两期情形中消费和资产份额的解与多期情形类似（即与期限无关）。消费财富比率和资产份额都独立于财富水平，但都依赖于预期资产收益。对对数效用函数而言，消费财富比率独立于财富水平（与幂效用函数情形相同），但也独立于预期收益。

* 说明SDF模型的一阶条件 $E\{R_{i,t+1}M_{t+1}\}$ 是如何从多期投资组合——消费优化问题中推导出来的。

15.1 两期模型

模型假设个体追求效用最大化，效用只取决于今天和明天的消费。个体选择最优消费数量 C_t 和 C_{t+1}，选择风险资产和无风险资产的持有数量。

我们案例详细地描述这个两期模型，因为两期模型的求解技巧也适用于多期模型（只有非常细小的改动）。求解技巧涉及后向递归，也就是说，我们先求出模型在最后一期 $t+1$ 的解，然后求解模型在第1期（即今天）的解。这种求解技巧使我们能看到多期框架中价值函数和贝尔曼（Bellman）方程的作用。正如我们即将看到的那样，即使在两期模型中，要得到消费和资产持有量的简单闭型解（closed form solution）都会非常困难。正如第14章有关CCAPM模型的讨论那样，这就是为什么投资组合－消费模型经常用某种闭型的欧拉方程或一阶条件来进行检验的原因。当然，人们始终可以用数字技巧来求解消费和资产持有量的一阶条件。但是，在数字技巧给定的条件下，我们始终可以改变外生变量或状态变量（如财富水平），度量这些变量对均衡消费和资产持有量的影响。

虽然两期模型是一个均衡模型，但它不是一般均衡模型，因为模型没有对资产收益变化的

基本来源（如生产率和实际投资）——也就是说，经济的供给方面——进行建模。

为了有助于模型的经济解释，我们假设效用函数为幂效用函数（$\gamma = 1$ 时为对数效用函数）。使用这类特殊效用函数的两个关键结论是：消费与财富成比例，资产份额与财富独立。

模型

个体选择消费（C_t，C_{t+1}）和资产份额 α_i 来使两期（时间可分离效用函数）效用达到最大：

$$\max U = U(C_t) + \theta E_t U(C_{t+1}) \qquad 0 < \theta < 1 \tag{1}$$

这里的 E_t 是期望算子。（尽管有些傲慢，我们还是用左边的 U 来表示“总效用”（我们本应该在这里引入另外的符号来避免将 t 期和 $t+1$ 期的效用函数 $U(C_t)$ 和 $U(C_{t+1})$ 相混淆）。但是，我们认为这里不会产生任何混淆。）因为不存在劳动收入，所以预算约束为

$$W_{t+1} = R_{t+1}(W_t - C_t) \tag{2}$$

其中，W_t 是实际财富的初始禀赋，R_{t+1} 是资产组合的（实际）收益（为了使记号更加简洁，我们不再用 R^*_{t+1} 表示总收益）。因此，

$$R = \sum_{i=1}^{n} \alpha_i R_i \qquad \left(\sum_{i=1}^{n} \alpha_i = 1\right) \tag{3}$$

α_i 等于财富中第 i 种资产的比例。假设无风险资产（第种资产）的收益为 R_f，于是 $m = n-1$，$\alpha_n = 1 - \sum_{i=1}^{m} \alpha_i$。因此，一旦我们求得了最优值 α_i（$i = 1, 2, \cdots, m$），那么应用上述恒等式就可以求出无风险资产的最优份额。投资组合的收益为

$$\begin{aligned} R &= \alpha_n R_f + \sum_{i=1}^{n} \alpha_i R_i \\ &= \left(1 - \sum_{i=1}^{m} \alpha_i\right) R_f + \sum_{i=1}^{n} \alpha_i R_i \\ &= R_f + \sum_{i=1}^{n} \alpha_i (R_i - R_f) \end{aligned} \tag{4}$$

我们已将约束条件 $\sum_{i=1}^{m} \alpha_i = 1$ 代入到方程（4）中，所以 α_i（$i = 1, 2, \cdots, m$）与有关的系数不存在任何约束。为了求解动态规划问题，我们需要一个终点条件。（不失一般性）假设遗产的数量为零，从而

$$W_{t+2} = 0 \qquad (\text{于是 } C_{t+1} = W_{t+1}) \tag{5}$$

最大值函数（1）可以写为

$$\begin{aligned} \max U &= U(C_t) + \theta U(W_{t+1}) \\ &= U(C_t) + \theta E_t U(R_{t+1}[W_t - C_t]) \end{aligned} \tag{6}$$

我们已经将预算约束（2）代入到（6），这个小“技巧”意味着我们无需用拉格朗日乘子就能对模型进行求解。注意，资产组合的收益 R_{t+1} 是 α_i 的函数。假设效用函数为幂函数

$$U(C) = C^{1-\gamma}/(1-\gamma) \qquad (0 < \gamma < 1) \tag{7a}$$

$$U'(C) = C^{-\gamma} \tag{7b}$$

因为 $C_{t+1} = W_{t+1}$，所以要注意

$$U'(C_{t+1}) = U'(W_{t+1}) = W_{t+1}^{-\gamma} \tag{7c}$$

我们已经知道，最大化问题一般形式的解（个体的“选择变量”C_t 和 α_{it}（$i = 1, 2, \cdots, m$））必须取决于外生变量或状态变量（也就是说 W_t 和 $E_t R_{i,t+1}$），当然也包括诸如 γ 和 θ 等参数。因此，一般形式的闭型解为：

$$C_t = h(W_t, Ef(R_{i,t+1}), R_f, \theta, \gamma)$$

$$\alpha_{it} = g(W_t, Ef(R_{i,t+1}), R_f, \theta, \gamma)$$

其中，f 和 g（可能的）是非线性函数（通常很难或几乎不可能获得解析解）。

最优消费：幂效用函数

（6）关于 C_t 的一阶条件为

$$\frac{\partial U}{\partial C_t}=0=U'(C_t)+\theta E_t\left(\frac{\partial U}{\partial W_{t+1}}\right)\left(\frac{\partial W_{t+1}}{\partial C_t}\right) \tag{8a}$$

假设效用函数为幂函数并使用（7b）和（2）就得到

$$0=C_t^{-\gamma}+\theta E_t\{W_{t+1}^{-\gamma}(-R_{t+1})\} \tag{8b}$$

将（2）中的 W_{t+1} 代入并进行重新排列，

$$\left(\frac{W_t-C_t}{C_t}\right)=[\theta E_t\{R_{t+1}^{(1-\gamma)}\}]^{1/\gamma}=E_t[f(R_{t+1}(\alpha_i),\gamma,\theta)] \tag{9}$$

其中，$f(\cdot)$ 是 R_{t+1} 的非线性函数，R_{t+1} 本身也是资产份额 α_i 和个体资产收益（$R_{i,t+1}$，R_{ft}）的函数。重新整理后就得到

$$C_t^*=\frac{1}{1+E_tf(R_{t+1}(\alpha_{it}),\theta\gamma)}W_t \tag{10}$$

因此，最优消费与财富成比例且取决于投资机会集（也就是说，所有风险资产和无风险资产的预期收益构成的集合）。消费财富比率取决于 $E_t[f(R_{t+1},R_f,\theta,\gamma)]$，因此预期投资组合的收益 R_{t+1} 对当前消费的影响依赖于收入效应和替代效应（见后）——更高的预期收益倾向于降低当前消费（替代效应）并增加下一期的财富（收入效应）。正如我们即将看到的那样，对 $\gamma=1$ 的特殊情形而言，两种效应可以相互抵消，消费独立于预期收益。注意，（10）不是 C_t 的完全闭型解，因为 R_{t+1} 依赖于决策变量 α_i。

最优资产份额

α_i 的一阶条件为

$$\begin{aligned}\frac{\partial U}{\partial \alpha_i}&=0\\&=\theta E_t\left(\frac{\partial U}{\partial W_{t+1}}\right)\left(\frac{\partial W_{t+1}}{\partial \alpha_i}\right)\\&\equiv\theta E_t\left\{U_W\frac{\partial W_{t+1}}{\partial \alpha_i}\right\}\ (i=1,2,\cdots,n)\end{aligned} \tag{11}$$

使用（2）、（4）和 $U(W_{t+1})$ 的幂函数形式就得到，

$$0=\theta E_t\{W_{t+1}^{-\gamma}(W_t-C_t)(R_i-R_f)\} \tag{12}$$

将（2）中的 W_{t+1} 代入，

$$\begin{aligned}0&=\theta E_t\{[(W_t-C_t)R_{t+1}]^{-\gamma}[(W_t-C_t)(R_{i,t+1}-R_f)]\}\\0&=E_t\{R_{t+1}^{-\gamma}(R_{i,t+1}-R_f)\}\end{aligned} \tag{13}$$

因为 $\theta(W_t-C_t)^{1-\gamma}$ 可以从数学期望中拿出来（因为它在时刻 t 是已知的），所以可以将其“消掉”。记住，R_{t+1} 依赖于 α_i，所以代入（4）就得到

$$0=E_t\left\{\left[R_f+\sum_{i=1}^{m}\alpha_{it}(R_{i,t+1}-R_f)\right]^{-\gamma}(R_{i,t+1}-R_f)\right\} \tag{14a}$$

方程（14a）包含 m 个方程，求解方程组就得到 α_{it} 的最优值。（14a）的解可以表示为

$$\alpha_{it}^*=g(E_t[f(R_{k,t+1})],R_{ft},\gamma)\qquad k=1,2,\cdots,m \tag{14b}$$

（14）比初次看上去更加有趣。假设只存在一种风险资产 i，从而（14）是单一随机变量 R_i 非线性函数的条件期望。一旦我们知道 R_i 的概率密度函数，我们能计算出期望值并得到 α_{it} 的解。当我们有 m 种风险资产时，任意 α_{it} 的最优解依赖于 m 种风险资产的（单期）条件期望收益。因此，我们需要知道单期收益的多元分布。一个富有启发性的想法是，方程（14）可以用来求解 $\alpha_{it}=g(f(E_tR_k),R_f,\gamma)$（$k=1,2,\cdots,m$）。由（14）所得到的关键结果是风险资产

的持有比例与当前财富独立（从而与前面的储蓄－消费决策独立）。因此，一旦有了解 α_{it}^*，我们就能利用方程（10）来求解 C_t^*/W_t，求解时可以把 C_t^*/W_t 当成所有资产预期收益的非线性函数。这种“可分”性质根源于等弹性效用函数，如幂效用函数（及 $\gamma=1$ 时的对数效用函数）。α_{it}^* 独立于当前财富的事实在对于在对个体资产进行加总时变得非常有用——尽管我们在此不会停留太多的时间。然而要注意的是，一般而言，一阶条件意味着 α_{it}^* 同时取决于财富和期望收益 ER_i，因此（10）和（14b）不得不同时进行求解（如二次效用函数的情形）。

解：幂效用函数情形

* 资产份额只取决于期望收益（和 γ），它与财富独立。

* 消费财富比率只依赖于期望收益，它与财富独立（因为 α_i 与财富独立，参见方程（14b））。

值得注意的是，在方程（6）中最大值函数为

$$U(W_t)=U(C_t)+\theta E_t U(W_{t+1}) \tag{14}$$

最大值函数可以解释为在 t 时刻的确定性函数，总效用 $U(W_t)$ 依赖于外生状态变量 W_t。方程（6）也是一个递归方程。从 $t+1$ 时刻起向后求解，我们就找到了最优值 C_t^* 和 α_{it}^*，于是使用（2）就可以得到 $W_{t+1}=R_{t+1}(W_t-C_t)$。因此，我们也能计算出（6）中效用函数的最优值 $U(W_t)$。递归关系在多期框架中非常重要，因为后者可以视为一系列（与（6）一样）单期问题。在多期情形中，将已经得到的 $U(W_t)$ 代入 $E_{t-1}(W_t)$ 的右边，然后根据一阶条件来计算 C_{t-1} 和 α_{t-1} 的最优值，最后就得到 $U(W_{t-1})$ 的值。在多期情形中，我们一直进行后向递归，直到我们能确定 $t=0$ 时的 C_0 和 α_{0i}。这就是随机动态规划中的贝尔曼递归最优求解方法。

解：$\gamma=1$ 时的对数效用函数情形

* 注意，当 $\gamma=1$ 时，方程（14）中 α_i 一阶条件的一般形式不会发生改变，因此最优值 α_i^* 仍然与财富水平独立。

* 消费不仅与财富成比例，而且与投资机会集（如预期收益）独立。当 $\gamma=1$ 时，由一阶条件（9）可知

$$C_t=\frac{W_t}{1+\theta} \tag{15}$$

预期收益变化的收入效应和替代效应在这里相互抵消，所以消费仍然没有变化。

数学期望的计算

有些读者在计算上述非线性方程中的数学期望时存在一些困惑。从原理上讲，计算应该很简单，因为当概率密度函数为 $f(x)$ 时连续函数 $g(x)$（如 $g(x)=x^2$）的数学期望为

$$E_t[g(x)]=\int_0^{\infty} g(x)f(x_t \mid x_{t-1},x_{t-2},\cdots)\,\mathrm{d}x \tag{16}$$

其中，$x=R_i$ 为总收益，积分区间为 $[0,+\infty)$。要么用解析方法要么用数值方法（更为一般的方法）来计算积分值。例如，可以假设 $f(x)$ 为正态密度函数。密度函数 $f(\cdot)$ 为条件密度，因为 t 时期的收益可能取决于 $t-1$ 期或更早时期的收益（即时变收益或可预测收益）。在 *iid* 收益的特殊情形中，条件分布与非条件分布相同。

为了看清方程（16）给出的是一个已知的“数值”，考虑一个简单的离散情形，总收益只有两种可能的结果 λ 和 $1/\lambda$（λ 是已知的，如 $\lambda=1.1$），且

$$\text{Prob}\{R_i=\lambda\}=1/2=\text{Prob}\{R_i=1/\lambda\} \qquad (\lambda>1) \tag{17}$$

如果 $g(R)=R$，那么（举个例子）

$$E[g(R)]=ER=(1/2)\lambda+(1/2)\lambda^{-1}=1.0045 \tag{18a}$$

如果 $g(R)=R^2$，那么

$$E\,[g\,(R)] = ER^2 = \frac{\lambda^2}{2} + \frac{\lambda^2}{2} = 1.0182 \qquad (18b)$$

现在，我们已经知道两期消费投资组合问题的求解技巧。对任意个体而言这些一阶条件都成立，因此我们就得到了每个人的最优解 C_i^* 和 α_{ii}^*。

寻找均衡收益的 CCAPM/SDF 模型

为了从个体的一阶条件出发来得到整个市场的一阶条件，我们需要资产供给方面的一些假设，要借助于“代表性投资者”的假设（即效用函数相同）。然后，我们分析均衡资产收益的行为。

在前面的章节中，我们的注意力集中在消费和资产份额的闭型解上，但模型必须有 CCAPM/SDF 模型的“隐含”一阶条件。事实确实如此，用拉格朗日乘数法来对最大化问题进行设定就可以看清这一点。我们有

$$\max U = U\,(C_t) + \theta E_t\,\{U\,(C_{t+1})\} \qquad (19)$$

预算约束为

$$C_{t+1} = W_{t+1} = R_{t+1}\,(W_t - C_t)$$

$$R_{t+1} = \sum_{i=1}^{n} \alpha_i R_{i,t+1} \qquad (20)$$

其中的一种资产可能是无风险资产，预算约束为

$$\sum_{i=1}^{n} \alpha_i = 1 \qquad (21)$$

拉格朗日函数为

$$U\,(C_t) + \theta E_t\,\{U\,(C_{t+1})\} + \lambda\,(1 - \sum_{i=1}^{n} \alpha_i) \qquad (22)$$

使用（20）并就（22）对 C_t 和 α_i 求导，一阶条件为

$$U'\,(C_t) = \theta E_t\,\{U'\,(C_{t+1})\,R_{t+1}\} \equiv \theta E_t\left\{U'\,(C_{t+1})\sum_{i=1}^{n} \alpha_i R_{i,t+1}\right\} \qquad (23)$$

$$\theta E_t\,\{U'\,(C_{t+1})\,R_{i,t+1}\} - \lambda/\,(W_t - C_t) = 0 \quad (i = 1,\,2,\,\cdots,\,n) \qquad (24)$$

从（24）可以看出，投资者在资产 i 和 j 之间一直进行选择直到 $t+1$ 时的期望边际效用相等

$$E_t\,\{U'\,(C_{t+1})\,R_{i,t+1}\} = E_t\,\{U'\,(C_{t+1})\,R_{j,t+1}\} \qquad (25)$$

（24）乘以 α_i，从 1 到 n 进行求和，然后代入（23）就得到

$$U'\,(C_t) = \sum_{i=1}^{n} \alpha_i\,[\lambda/\,(W_t - C_t)] = \lambda/\,(W_t - C_t) \qquad (26)$$

将（26）代入（24）就得到与 SDF 模型类似的一阶条件

$$\theta E_t\left\{\frac{U'\,(C_{t+1})}{U'\,(C_t)} R_{i,t+1}\right\} = E_t\,\{M_{t+1} R_{i,t+1}\} = 1 \quad (i = 1,\,2,\,\cdots,\,n)$$

上述方程包含无风险利率的一阶条件 $R_f E_t\,(M_{t+1}) = 1$。这种方法与前面“扰动方法”（见 11 章）之间的唯一差异在意，我们已经引入了显式预算约束，并直接对带约束的最大化问题进行求解。因此，对这个消费投资组合模型而言，11 章的所有结论都成立。

也许值得一提的是，（非随机）外生劳动收入 Y_t 的引入本应该导致如下预算约束

$$C_{t+1} \equiv W_{t+1} = R_{t+1}\,(W_t + Y_t - C_t)$$

但一阶条件的核心方程保持不变（要用 $W_t + Y_t$ 来替换 W_t）。

15.2 多期模型

在资产市场中，代表性投资者扮演着价格接受者的角色，选择当前消费和资产份额来使终

生效用达到最大。正如我们即将看到的那样，借助价值函数的概念，多期问题可以转化为一系列两期问题。价值函数是一种递归关系，可以从最后时刻开始进行向后递归求解，在随机动态规划文献中，经常称价值函数为贝尔曼方程。因此，在 t 时刻，为了计算今天的最优值，你必须从 T 到 t “向后” 求解所有的最优值。然而，在 t 时刻已经选择了最优值 C_t^* 和 C_{t+1}^*后，但 t 到 $t+1$ 之间仍会出现新信息。因此，在 $t+1$ 时刻，我们必须重新计算新的最优值 $\widetilde{C}_{t+1}^*$，这个最优值一般不等于在 t 时刻为 $t+1$ 时刻所选择的最优值（即 $C_{t+1}^* \neq \widetilde{C}_{t+1}^*$）。跨期问题在时刻一般的闭型解具有如下形式：

$$C_t = f\ (W_t,\ E_tR_{t+j},\ \gamma,\ \theta,\ T-t) \tag{27a}$$

$$\alpha_{it} = g\ (W_t,\ E_tR_{t+j},\ \gamma,\ \theta,\ T-t) \tag{27b}$$

这里的 γ 是跨期效用函数的参数。闭型解通常无法得到，但对幂效用函数和对数效用函数（$\gamma=1$）之类的时间可分离效用函数而言，我们在获得闭型解方向上取得了一些进展。这些跨期模型在连续时间情形中有时更容易处理，但是，为了保持一致（可能要增加一些理想条件），我们只讨论离散情形。

偏好和预算约束

假设个体追求时间可分离效用函数的最大化，这种类型的效用函数只取决于当前消费 C_0 和所有未来消费（C_1，C_2，…）。当前时刻为 $t=0$，

$$U = E_0 \sum_{t=0}^{T-1} U\ (C_t,\ t)\ + B\ (W_T,\ T) \tag{28}$$

这里的 B（·）是 T 时刻的遗赠函数。我们继续使用幂效用函数，因为尽管这是一种特殊情形，确实省去了繁琐的代数运算，还能使我们窥视一阶条件内部的 “黑盒子”。对幂效用函数而言，我们有

$$U\ (C_t,\ t)\ = \theta^t C_t^{1-\gamma} /\ (1-\gamma) \tag{29a}$$

$$B\ (W_T,\ T)\ = \theta^T C_T^{1-\gamma} /\ (1-\gamma) \tag{29b}$$

为了简化符号，我们将可分离的时间变量省略，例如 $U\ (C_t,\ t)\ \equiv U\ (C_t)$。预算约束为

$$W_{t+1} = R_{t+1}\ (W_t - C_t) \tag{30}$$

资产组合的收益为

$$R_{t+1} = \sum_{i=1}^{m} \alpha_{it}\ (R_{i,t+1} - R_{ft})\ + R_{ft} \tag{31}$$

α_{it}是待定的资产份额，R_i 是 $m=n-1$ 种风险资产的收益，无风险资产的份额为 $\alpha_n = 1 - \sum_{i=1}^{m}\alpha_i$。求解方法与两期情形中的上述方法很接近，尽管代数运算要更为详细、更为复杂一些（因此所有求和公式和数学期望算子都成立）。注意，尽管我们给出的解包含 m 个 α_{it}值，如果只有一种风险资产（和一种无风险资产），那么计算过程会简洁得多。

价值函数

为了求解这些多期问题，我们引入价值函数 $J\ (W_T)$（有时称为财富引致效用函数），价值函数的定义如下：

$$J\ (W_T)\ = \max E_t \left\{ \sum_{s=t}^{T-1} U\ (C_s)\ + B\ (W_T) \right\} \tag{32}$$

注意，$J\ (W_T)$ 是 t 时刻效用函数的最大值。因此，最优决策是从 $t \geqslant 0$ 的任意时刻开始向前进行的。价值函数允许我们将多期问题表示为一系列两期问题。于是，

$$J\ (W_t)\ = B\ (W_t) \tag{33}$$

在 $T-1$ 时刻，个体投资者选择 C_{T-1}和 α_{T-1}来对函数进行最大化

$$J\ (W_{T-1})\ = \max_{C_{T-1},\alpha_{T-1}} \{U\ (C_{T-1})\ + E_{T-1}B\ (W_T)\} \tag{34}$$

将（30）代入，由（31）可以注意到 R_{t+1} 是 α_{it} 的一个函数，（34）就变成

$$J(W_{T-1})=\max_{C_{T-1},\alpha_{T-1}}\{U(C_{T-1})+E_{T-1}\{B(R_T(\alpha_{i,T-1})(W_{T-1}-C_{T-1}))\}\} \tag{35}$$

我们就得到了一个类似的两期问题。C_{T-1} 的一阶条件为

$$0=U'(C_{T-1})+E_{T-1}\{B_W(W_t)(\partial W_T/\partial C_{T-1})\} \tag{36}$$

对幂效用函数而言有 $B_W=\partial B/\partial W_T=\theta^T W_T^{-\gamma}$，由（30）可知

$$\partial W_T/\partial C_{T-1}=-R_T \tag{37}$$

因此，（36）通常可以表示为

$$0=U'(C_{T-1})-E_{T-1}\{B_W(W_T)R_T(\alpha_i)\}) \tag{38a}$$

$$=U'(C_{T-1})-E_{T-1}\left\{B_W(W_T)\left[\sum_{i=1}^{m}\alpha_{i,T-1}(R_{i,T}-R_f)+R_f\right]\right\} \tag{38b}$$

对（34）求微分并使用（30）和（31）就得到 α_i 的一阶条件

$$0=E_{T-1}\{B_W(W_T)(R_{i,T}-R_f)\}\qquad(i=1,2,3,\cdots,m) \tag{39}$$

（39）乘以 α_i 然后求和（$i=1,2,3,\cdots,m$）就得到

$$E_{T-1}\left\{\sum_{i=1}^{m}\alpha_i R_{i,T}B_W(W_T)\right\}=R_f E_{T-1}\{B_W(W_T)\} \tag{40}$$

或

$$0=E_{T-1}\left\{B_W(W_T)\sum_{i=1}^{m}\alpha_i(R_{i,T}-R_f)\right\} \tag{41}$$

将（41）代入消费的一阶条件（38b）就得到另外一种表示

$$U'(C_{T-1})\ R_f E_{T-1}\{B_W(W_T)\} \tag{42}$$

一阶条件（38b）和 m 个一阶条件（39）是关于 $m+1$ 个未知数（C_{T-1} 的值和 $\alpha_{i,T-1}$ 的 m 个值）的 $m+1$ 个方程。求解这个方程组就可以得到 $T-1$ 时刻的最优资产份额和消费。我们针对幂效用函数的特殊情形来进行求解

$$B_W(W_T)=\frac{\partial}{\partial W_t}\theta^T W_T^{1-\alpha}/(1-\alpha)=\theta^T W_T^{-\gamma} \tag{43a}$$

$$U'(C_{T-1})=\theta^{T-1}C_{T-1}^{-\gamma} \tag{43b}$$

将（43a）和（43b）代入一阶条件（38a），用（30）和（31）替代 W_T，经过一些运算之后我们就得到

$$\theta^{T-1}C_{T-1}^{-\gamma}=\theta^T(W_{T-1}-C_{T-1})^{-\gamma}E_{T-1}\{R_T^{1-\gamma}(\alpha_i)\} \tag{44}$$

或

$$\frac{C_{T-1}^{*}}{W_{T-1}}=\frac{1}{1+[\theta E_{T-1}\{R_T^{1-\gamma}(\alpha_i)\}]^{1/\gamma}} \tag{45}$$

$R_T(\alpha_i)$ 意味着 R_T 是资产份额的函数。与幂效用函数情形的两期问题一样，我们发现最优消费与财富成比例，也依赖于预期收益。注意，在 $\gamma=1$ 的对数效用函数中，消费与财富成比例但与变化的投资机会（即资产的期望收益）独立，这与两期模型相同。使用幂效用函数，我们转到资产份额一阶条件（40）上去

$$E_{T-1}\{R_{i,T}B_W(W_T)\}=E_{T-1}\{R_f B_W(W_T)\} \tag{46a}$$

使用（43a）就得到

$$E_{T-1}\{R_{i,T}W_T^{-\gamma}\}=E_{T-1}\{R_f W_T^{-\gamma}\} \tag{46b}$$

现在，将（30）中的 $W_T=(W_{T-1}-C_{T-1})R_T$ 和（31）中的 R_T 代入并消去（$W_{T-1}-C_{T-1}$）后就得到

$$E_{T-1}\left[(R_i-R_f)\left(\sum_{i=1}^{m}\alpha_{i,T-1}(R_{iT}-R_f)+R_f\right)^{-\gamma}\right]=0\quad(i=1,2,\cdots,m) \tag{47}$$

方程（47）是关于了 m 个未知数 α_i 的个方程，一旦我们知道了收益的密度函数，我们就可以求解

$$\alpha_{i,T-1}^{*}=g(E_{T-1}R_i,\ E_{T-1}R_j,\ \cdots,\ R_f) \tag{48}$$

与两期问题中的情形一样，α_i^* 独立于财富（因此所有具有幂效用函数的投资者都持有相同比例的资产）。当然，幂效用是一种非常特殊的情形，一般而言，$\alpha_{i,T-1}$ 也是财富的函数。资产比例也取决于期望收益，但是，如果收益独立同分布，那么无论投资者的计划时期 $T-t$ 多长，$\alpha_{i,t}$ 都相同。我们也经常称这种特殊情形为短视行为，因为在多期模型中，最优资产比例与单期投资者的最优资产比例相同。

根据（47）解出 $\alpha_{i,T-1}^*$，我们可以将其代入（45）来求解 C_{T-1}^*/W_{T-1}。与幂效用函数的两期情形一样，我们不需要对 C_{T-1}^* 和 α_{T-1}^* 同时求解。

无限递归

因为我们已经得到了 C_{T-1}^* 和 $\alpha_{i,T-1}^*$ 的最优值，所以由（34）就可以计算出 $J(W_{T-1})$ 的最优值。使用价值函数（34）我们可以将 $T-2$ 时的最优化问题设定为

$$J(W_{T-2}) = \max_{C_{T-2},\alpha_{T-2}} E_{T-2}\{U(C_{T-2}) + U(C_{T-1}^*) + B(W_T)\} \tag{49}$$

在 $T-2$ 时刻，W_{T-2} 是已知的，根据最优化原理，在 $T-2$ 时刻我们可以选择 C_{T-2} 和 $\alpha_{i,T-2}$，在 $T-2$ 时刻的这一结果给定时，$T-1$ 时的决策必须是最优的。此外，使用迭代期望公式 $E_{T-2}E_{T-1}(X)=E_{T-2}(X)$，我们可以将（49）写成

$$J(W_{T-2}) = \max\{U(C_{T-2}) + E_{T-2}(\max_{C_{T-1},\alpha_{i,T-1}} E_{T-2}[U(C_{T-1}^*) + B(W_T)]\} \tag{50}$$

$$J(W_{T-2}) = \max\{U(C_{T-2}) + E_{T-2}[J(W_{T-1})]\} \tag{51}$$

$J(W_{T-1})$ 可以根据 $T-1$ 时刻所得到的最优值 C_{T-1}^* 和 $W_{i,T-1}^*$ 来进行计算。从直观来看，（51）中的效用函数依赖于当期消费 C_{T-2} 和下期财富 W_{T-1}。如果我们在 $T-2$ 时刻消费得更多，那么下期财富就会减少，因此这是一个权衡。

方程（51）是关于 $J(\cdot)$ 的一个递归关系式，（51）的求解与（34）有相同的形式，我们已经注意到这一点。因此，一般而言，对任意时刻 $t=0, 1, 2, \cdots, T-1$，与（42）和（39）类似的最优条件（用 J_W 替代 B_W）为

$$U'(C_t) = R_{ft}E_tJ_W(W_{t+1}) = J_W(W_t) \tag{52a}$$

$$E_t\{R_{i,t+1}J_W(W_{t+1})\} = R_{ft}E_t\{J_W(W_{t+1})\} \quad (i=1, 2, \cdots, m) \tag{52b}$$

在上述结论中，我们没有明确推导的部分是在（52a）中的包络条件

$$U'(C_t) = J_W(W_t) \tag{53}$$

消费一阶条件的这种表示意味着，沿着最优路径，储蓄 1 美元的边际效用等于消费 1 美元的边际效用。对价值函数（51）求全微分然后用一阶条件来进行替换就得到包络条件，附录 I 进行了推导。

（52a）和（52b）中的一阶条件看上去非常简洁，但对任意给定的效用函数而言仍然需要进行求解，闭型解只对容许效用函数的一个子集才是可能的。一般而言，任意时刻 t 的最优资产比例和消费都依赖于财富、t 到 $T-1$ 的预期收益和效用函数的参数。要求出完整的闭型解往往是非常困难的或不可能推导出来。因此，模型通常是在一阶条件/欧拉方程的基础上用回归技巧来进行实证检验，正如我们在前面章节中所看到的那样。

即使在对数效用函数情形中，（47）或（52b）中最优权重 α_i 的解在只有一种风险资产时也不是很直接，因为我们必须计算收益的非线性函数的数学期望。然而，C_t 的最优解（见附录 II）为

$$C_t = \frac{1-\theta}{1-\theta^{T-t+1}}W_t \tag{53}$$

因此，具有对数效用的个体计划在有生之年（t 到 T）“匀滑地”消费他们的财富。

如果收益独立同分布且效用函数为幂效用函数，那么资产需求与投资年限独立（这就是所谓的近视行为），因为在投资者具有前瞻性的多期模型中，最优资产比例与单期投资者的最优资产比例相同。

在思维上跟上上述推导后，读者就可以学习 Samulson（1969）了。但是，随机动态规划这种求解技巧和价值函数背后的基本思想仍然广为使用，因为它们有助于寻找很多跨期随机模型的可行试验解。然而，在许多特殊情形中，数值解法仍然是当前经常使用的方法（参见 Barberis、Huang 和 Santos（2001））。例如，当效用函数为对数效用函数时，我们知道消费的直观解为（53）。与此相对照的是，假设 W_t 是已知的模型输入，数值方法将给出 C_t 的唯一值，由此可以推导出（53）中的线性关系（也就是说，我们可以画出 W_t 的各个输入值与相应数值最优解 C_t 之间的对应图像）。因此，在更为复杂的情形中，我们可以使用包含数值模拟的可计算一般均衡模型。事实上，一旦知道了状态变量 W_t 和期望收益的预期值（未来各期），我们就可以获得当前和未来各期的最优消费路径 C_t，C_{t-1}，…，C_T 和最优资产份额。如果知道任意时刻 t 的最优值 C_t^* 和 α_{it}^*，由（28）或递归地使用贝尔曼方程，我们也就知道了一生的效用。

这些类型的模型在经济学中非常盛行，尤其是在经济行为具有跨期成分和约束的情形中非常盛行。如“最优”税率的设计，养老金和福利政策要求政府部门为工作或退休的人设定税率和福利利率。约束条件可能包括终生预算约束（包括福利收入和国家养老金支付）、最低福利水平（在任何时期）、有关劳动供给和资产收益的假设。使终生效用（取决于消费和闲暇）达到最大的税率和福利水平是可以计算的，可以对各种政策情景（如不同的收入税或不同的风险规避程度）中的终生效用进行比较。因此，理解这类模型背后的直觉和技巧问题非常重要，我们以消费－投资组合模型作为讨论的开始，后面将对此进行深入讨论。

期望值必须用数值方法来进行计算。在上述模型中，唯一的“外生”随机变量就是资产收益。在数学期望的计算过程中，如果我们将“不确定的”劳动收入加入到预算约束中去，那么我们不得不涉及收益和劳动收入之间的相关性，这使得数值求解程序变得相当的复杂（Campbell 和 Vicera（1999））。

15.3 期望收益的 SDF 模型

通过两期模型，我们能考察跨期消费－投资组合选择模型的一阶条件，推导出用 SDF 形式表示的一阶条件。由价值函数导出的一阶条件的简洁形式在这里非常有用，它几乎可以直接给出我们所需要的解（因此，进行推导是值得的）。消费和资产投资份额的一阶条件为

$$U'(C_t)=J_W(W_t)=R_{ft}E_t\{J_W(W_{t+1})\} \tag{54}$$

$$E_t\{R_{i,t+1}J_W(W_{t+1})\}=R_{ft}E_t\{J_W(W_{t+1})\}\qquad(i=1,2,\cdots,m) \tag{55}$$

（54）中的包络条件对所有 t 都成立，因此

$$U'(C_{t+1})=J_W(W_{t+1}) \tag{56}$$

将（56）代入（54）并进行重新整理就得到无风险利率的 *SDF* 方程

$$R_{ft}E_t(M_{t+1})=1 \tag{57}$$

其中，$M_{t+1}=U'(C_{t+1})/U'(C_t)$ 是 SDF 或定价核。（注意，我们在这里将 θ“吸收到”M_{t+1} 中，因为 $U(C_t)=\theta^t f(C_t)$，$f(C_t)$ 是普通效用函数，如 $f(C_t)=\ln C_t$。）

对风险资产而言，我们也将（56）代入（55），应用（54）就得到

$$E_t\{R_{i,t+1}U'(c_{t+1})\}=U'(C_t)\qquad(i=1,2,\cdots,m) \tag{58}$$

$$E\{R_{i,t+1}M_{t+1}\}=1 \tag{59}$$

于是，我们在第 10 章和第 11 章详细讨论过的 SDF 模型与跨期消费－投资组合模型一致。值得注意的是，（58）本身是一个递归式，对任意投资期限都成立，因为

$$U'(C_t)=E_t\{R_{i,t+1}U'(C_{t+1})\}=E_t\{R_{i,t+1}R_{i,t+2}U'(C_{t+2})\} \tag{60}$$

于是

$$1=E_t\{R_{i,t+2}M_{t,t+2}\} \tag{61}$$

其中，$X_{t,t+2}=X_{t,t+1}X_{t+1,t+2}$，$X=R_i$ 或 M。

15.4 小结

* 具有时间可分离效用函数和一个状态变量（财富）的多期消费 - 投资组合问题可以使用价值函数/贝尔曼方程从 T 时刻开始“向后”递归求解。

* 一般而言，最优资产份额（风险资产和无风险资产）和最优消费财富比率依赖于财富、预期收益和风险规避程度。

* 在等弹性边际效用（即幂效用函数和对数效用函数）的特殊情形中，最优一阶条件可以独立地进行求解，先计算资产份额，然后计算消费财富比率。但是，应同时对由这两个一阶条件所组成的方程组进行求解（在任意时点上）。

* 对幂效用函数而言，最优资产份额和消费财富比率独立于财富，但依赖于预期收益。对 $\gamma=1$ 的对数效用函数而言，消费财富比率独立于预期收益（只依赖于 θ 和时间 $T-t$）。

* 仿射模型（即 CCAPM 模型）的一阶条件或欧拉方程能从消费 - 投资组合问题的一阶条件导出。

附录 I：消费 - 投资组合问题的包络条件

我们按照如下程序来导出包络条件。$T-1$ 时刻的价值函数（也称为 Bellman 方程）为

$$J(W_{T-1}) = \max_{C_{T-1},\alpha_{i,T-1}} \{U(C_{T-1}) + E_{T-1}B(W_T)\} \tag{A1a}$$

现在假设已经解出了最优值 $\{C_{T-1}, \alpha_{T-1}\}$，我们将这些值代回到价值函数中去。使用（A1a）中的最优值，$T-1$ 时刻的 Bellman 方程为

$$\begin{aligned} J(W_{T-1}) &= U(C_{T-1}) + E_{T-1}\{B(W_T)\} \\ &= U(C_{T-1}) + E_{T-1}B\Big\{\Big[\sum_{i=1}^{m}\alpha_i(R_i-R_f)+R_f\Big](W_{T-1}-C_{T-1})\Big\} \end{aligned} \tag{A1b}$$

因为 $\{C_{T-1}, \alpha_{T-1}\}$ 是最优值，所以我们可以将“max”省略掉。对 $J(W_{T-1})$ 针对 W_{T-1} 求全微分，注意 $dB(W_T) = (\partial B/\partial W_T)\,dW_T$，$dW_T$ 依赖于 W_{T-1}、$\alpha_{i,T-1}$ 和 C_{T-1}，因为

$$W_T = (W_{T-1}-C_{T-1})\,R_T(\alpha_{t,T-1}) \tag{A2}$$

$$\begin{aligned} J_{W(T-1)} &= U'(C_{T-1})\frac{\partial W_T}{\partial W_{T-1}} \\ &\quad + E_{T-1}\Big\{B_{W(T)} \times \Big[\frac{\partial W_T}{\partial W_{T-1}} + \frac{\partial W_T}{\partial C_{T-1}}\frac{\partial C_{T-1}}{\partial W_{T-1}} + \sum_{i=1}^{m}\frac{\partial W_T}{\partial \alpha_i}\frac{\partial \alpha_i}{\partial W_{T-1}}\Big]\Big\} \end{aligned} \tag{A3}$$

其中，$J_{W(T-1)} = \dfrac{\partial J(W_{T-1})}{\partial W_{T-1}}$。使用定义

$$W_T = \Big[\sum_{i=1}^{m}\alpha_{i,T-1}(R_{i,T}-R_f)+R_f\Big](W_{T-1}-C_{T-1}) \equiv R_T(W_{T-1}-C_{T-1}) \tag{A4}$$

来计算偏导数，我们有

$$\begin{aligned} J_{W(T-1)} &= U'(C_{T-1})\frac{\partial C_{T-1}}{\partial W_{T-1}} + E_{T-1}\Big\{B_{W(T)}\Big[R_T\Big(1-\frac{\partial C_{T-1}}{\partial W_{T-1}}\Big)\Big] \\ &\quad + B_{W(T)}(W_{T-1}-C_{T-1})\sum_{i=1}^{m}(R_{i,T}-R_f)\frac{\partial \alpha_i}{\partial W_{T-1}}\Big\} \end{aligned} \tag{A5}$$

这个表达式看起来有些糟糕，但是要注意的是，正文中已经推导出最优值 $\{C_{T-1}, \alpha_{T-1}\}$ 的一阶条件（38a）和（39），

$$E_{T-1}\{B_W R_T\} = U'(C_{T-1}) \tag{A6a}$$

$$E_{T-1}\ \{B_W\ (R_{i,T}-R_f)\}\ =0 \tag{A6b}$$

同时也要注意的是，$T-1$ 时刻的任何事情都是已知的，可以将相关项从数学期望中拿出来。因此，我们有

$$J_{W(T-1)}=U'\ (C_{T-1})\ \frac{\partial C_{T-1}}{\partial W_{T-1}}+E_{T-1}\left\{U'\ (C_{T-1})\ \left(1-\frac{\partial C_{T-1}}{\partial W_{T-1}}\right)+0\right\} \tag{A7}$$

化简后就得到

$$J_{W(T-1)}=U'\ (C_{T-1}) \tag{A8}$$

这就是我们所需要的包络条件，我们在正文中已经讨论过。从正文的（42）就可以看出，一阶条件的另外一个版本是 $U'\ (C_{T-1})\ =R_fE_{T-1}\ (B_W)$，因此我们也要注意

$$U'\ (C_{T-1})\ =R_fE_{T-1}\ \{B_{W(T)}\} \tag{A9}$$

完整的证明需要我们证明“（A8）和（A9）对任意时刻 t 而不仅仅是对 $T-1$ 时刻成立”，我们在此不再深入阐述。

附录 II：对数效用函数情形的解

通过求解最优消费和资产份额的多期问题来考察价值函数（贝尔曼方程）到底起多大作用是富有建设意义的。我们在最容易使用的情形中（对数效用函数情形）来进行这样的操作。由正文可知，最优值 C 和 α_i 的一阶条件可以独立求解，C/W 独立于期望收益和 α_i^*（后者大大简化了对 C/W 的求解，因为我们无法得到 α_i 的显式最优解）。

我们需要的步骤如下：

（ⅰ）求解 C^*_{T-1} 和 α^*_{T-1} 的一阶条件；

（ⅱ）将解（C^*_{T-1}，α^*_{T-1}）代入贝尔曼价值函数中去，求 $J\ (W_{T-1})$ 的最优值；

（ⅲ）使用 $T-1$ 时刻的一阶条件（包括 $J_W\ (W_{T-1})$ 中的项）来求解最优值（C^*_{T-2}，α^*_{T-2}），然后进行递归求解。

我们有

$$B\ (W_T)\ =\theta^T\ln W_T \tag{A1}$$

$$U\ (C_{T-1})\ =\theta^{T-1}\ln C_{T-1} \tag{A2}$$

$$R_T=\sum_{i=1}^{m}\alpha_i\ (R_{i,T}-R_f)\ +R_f \tag{A3}$$

$$W_T=\ (W_{T-1}-C_{T-1})\ R_T \tag{A4}$$

在 $T-1$ 时刻，消费的一阶条件（A6a）为

$$U_c\ (C^*_{T-1})\ =\theta^{T-1}C^{-1}_{T-1}=\theta^TE_{T-1}\ \{B_W\ (W_T)\ R_T\}\ =\theta^T\ (W_{T-1}-C_{T-1})^{-1} \tag{A5}$$

这意味着

$$C^*_{T-1}=W_{T-1}/\ (1+\theta) \tag{A6}$$

对对数效用函数而言，C^*_{T-1} 独立于 $E_{T-1}\ [R_T\ (\alpha_i)]$，这意味着为了求解 C^*_{T-1} 我们不用求解 α_i^*。然而，资产份额的一阶条件（A6b）为

$$\begin{aligned}0\ &=E_{T-1}\ \{\ (R_{i,T}-R_f)\ B_W\ (W_T)\}\\&=E_{T-1}\ \{\ (R_{i,T}-R_f)\ /R_T\ (\alpha_i^*)\}\qquad (i=1,\ 2,\ \cdots,\ m)\end{aligned} \tag{A7}$$

我们在这里使用了（A1）和（A4），并将 $\theta^T/\ (W_{T-1}-C_{T-1})$ 中非随机项消掉了。

在求得（A7）中数学期望之后就可以对 m 个方程针对 m 个未知数 $\alpha_i^*\ (T-1)$ 进行求解，所得到的解独立于财富。

在最优化值 $\{C_{T-1}^{*},\ \alpha_{T-1}^{*}\}$ 处求 $J(W_{T-1})$

现在，我们来计算 $T-1$ 时刻的价值函数

$$J(W_{T-1}) = \underbrace{\max}_{C_{T-1},\omega_{T-1}} \{U(C_{T-1}) + E_{T-1}[B(W_T)]\}$$

$$= \underbrace{\max}_{C_{T-1},\omega_{T-1}} \{\theta^{T-1}\ln C_{T-1} + E_{T-1}(\theta^T \ln W_T)\} \tag{A8}$$

代入 C_{T-1}^{*} 和 $\alpha_{i,T-1}^{*}$ 并将表达式中的“max”去掉，

$$J(W_{T-1}) = \theta^{T-1}\ln[W_{T-1}/(1+\theta)] + \theta^T E_{T-1}\{\ln[(W_{T-1}-C_{T-1}^{*})R_T]\} \tag{A9}$$

再次将 $C_{T-1}^{*}=W_{T-1}/(1+\theta)$ 代入（A9），

$$J(W_{T-1}) = \theta^{T-1}(1+\theta)\ln W_{T-1} + k(\theta) + \theta^T E_{T-1}[\ln R_T(\alpha_{i,T-1}^{*})] \tag{A10}$$

其中，$k(\theta) = -\theta^{T-1}\ln(1+\theta) + \theta^T\ln(\theta/(1+\theta))$。

要记住的一个重要方面是，由一阶条件可知 $R_T(\alpha_{i,T-1}^{*})$ 独立于 W_{T-1}，因此

$$J_W(W_{T-1}) = \theta^{T-1}(1+\theta)/W_{T-1} \tag{A11}$$

$T-2$ 时刻的一阶条件

在 $T-2$ 时刻，消费的一阶条件为

$$U_c(C_{T-2}^{*}) = R_f E_{T-2} J_W(W_{T-1}) \tag{A12}$$

$$\theta^{T-2}/C_{T-2}^{*} = E_{T-2}\{\theta^{T-1}(1+\theta)R_f/W_{T-1}\}$$

$$= E_{T-2}\left\{\frac{\theta^{T-1}(1+\theta)R_f}{(W_{T-2}-C_{T-2}^{*})R_{T-1}}\right\} \tag{A13}$$

由（A13）可以看出，C_{T-2}^{*} 似乎依赖于 R_{T-1}，从而依赖于 $\alpha_{i,T-1}^{*}$。然而，我们可以证明，对对数效用函数目前而言，$R_f R_{T-2}\{1/R_{T-1}\}=1$，我们假设这个命题为真。由（A13）可以解得

$$C_{T-2}^{*} = \frac{W_{T-2}}{1+\theta+\theta^2} \tag{A14}$$

将递归程序推广到任意时刻 $t<T-1$，

$$C_t = \frac{W_t}{1+\theta+\theta^2+\theta^{T-t}} = \left[\frac{(1+\theta)}{1-\theta^{T-t+1}}\right]W_t \tag{A15}$$

余下要证明的是，对对数效用函数而言，

$$R_f E_{T-2}\{1/R_{T-1}\} = 1 \tag{A16}$$

$T-1$ 时刻消费一阶条件的一种表示是（见方程（42））

$$U_c(C_{T-1}^{*}) = E_{T-1}\{B_W(W_T)R_f\} \tag{A17a}$$

$$\theta^{T-1}/C_{T-1}^{*} = \theta^T R_f E_{T-1}\{1/(W_{T-1}-C_{T-1}^{*})R_T\} \tag{A17b}$$

代入 $C_{T-1}^{*}=W_{T-1}/(1+\theta)$ 并进行重新整理就得到

$$R_f E_{T-1}\{1/R_T(\alpha_i^{*})\} = 1 \tag{A18}$$

这对 $E_{T-2}\{1/R_{T-1}\}$ 也成立，因为一阶条件的形式与 $T-1$ 时刻相同。这就完成了证明的最后部分。

注意，我们没有直接解出 α_{it}^{*}（$t=T-1,\ T-2,\ \cdots$），因为这需要先求出数学期望的值。另外，对幂效用（$\gamma\neq1$）而言，C_t^{*} 的解依赖于期望收益和 $\alpha_{i,t}^{*}$。因此，尽管解仍然是闭型解，但求解过程还是更为复杂一些。

第16章

跨期资产配置：实证

目的

* 在跨期消费－投资组合模型中引入随机劳动收入。这就导致了风险资产的对冲需求，对冲需求取决于风险资产收益冲击和劳动收入冲击之间的相关系数。

* 考察风险资产投资份额依赖于投资年限的条件。

* 分析社会保障税投资于股票市场的含义。

* 考察最优资产配置中收益可预测性的定量影响。

* 说明股票收益预测方程中参数的不确定性如何影响最优资产的份额。

* 考察各种偏好假设对资产需求的含意。

16.1 退休与随机收入

跨期投资组合－消费模型对各种设定下（如考虑工作或退休、对国家或私人养老基金系统进行设计）消费和资产选择的分析而言非常重要。在这一节里，我们只是对这些模型进行总结，用到目前为止相对简单的方法来阐述模型之间的关系，在关键参数假设发生微小变化后对模型结论进行敏感性分析（特别地，我们引入随机劳动收入，由此可能导致风险资产的对冲需求）。

一个有趣的问题是，投资组合－消费模型是否与大众媒体所给出的标准建议（工作人员在风险资产上的投资比退休人员高）一致。另外，就工作人员而言，常用的“经验规则”是以60:40的比例投资于风险资产和无风险资产。

在我们的简单跨期模型中，效用函数是时间可分离的幂效用函数（无劳动收入），风险资产的份额（α）依赖于（条件）预期收益。Samuelson（1969）和 Merton（1969，1971）证明了如下结论：只有在非常特殊的情形中，投资期限（即 $T-t$）才对跨期消费－投资组合选择模型中风险资产的份额没有影响。这些特殊情形包括：幂效用函数、*iid* 资产收益、不存在劳动收入。单期解和多期解之间的等价性也经常被称为近视资产组合选择方法。因此，在这些假设条件下，资产份额在工作时期和退休时期中没有区别。

Bodie、Merton 和 Samuelson（1992）的跨期消费－投资组合模型（幂效用函数和 *iid* 收益）包括确定劳动收入。劳动收入是一种不可交易资产（即你不可能借入未来劳动收入，人力资本是未来收入的当前价值）。在这个模型中，确定的劳动收入所扮演的角色与无风险资产相同，并对无风险资产有“挤出效应”，导致风险资产的持有份额上升。因此，$\alpha^e > \alpha^r$，e 等于“就业”，r 等于“退休”。就业所得的确定性劳动收入意味着就业者可以将更多资金投入到风

险资产上去（与退休的情形相比，条件是假设劳动收入为零）。

如果劳动收入不确定且与风险收益无关，那么投资者对风险资产的偏好将弱于收入确定时的偏好（Viceira（2001））。但是，当随机收入与收益不相关时，高人力资本（即高的预期未来收入）家庭比低人力资本家庭持有更多的风险资产。

Viceira（2001）强调收入是随机的且风险资产收益同期相关时的投资组合 - 消费问题。现在，我们多了一个状态变量：劳动收入。这会产生风险资产的对冲需求（Merton（1971）在时变预期收益跨期模型中最先注意到这一点）。对冲需求根源于降低终生消费风险的愿望，这种风险又根源于收益和收入之间的相关性。如果收入冲击和收益冲击负相关，那么股票就是“合意的”，因为股票在收入低的时候能为投资者获得收益，从而有助于匀滑消费。

在 Viceira 的模型中，代表性投资者在预算约束下追求预期跨期幂效用的最大化

$$\max_{\{C_t,\alpha_t\}_0^\infty} E \sum_{t=0}^{\infty} \theta^t U\ (C_t) \tag{1a}$$

$$W_{t+1} = \ (W_t + Y_t - C_t)\ R_{p,t+1} \tag{1b}$$

$0<\theta<1$ 为折现因子，Y_t 是就业（劳动）收入，Y_t 在退休状态下为零。个体无法对劳动收入进行保险，你不可能针对未来收入签订和约。因为不存在劳动供给决策，所以收入是外生的。投资组合的收益为

$$R_{p,t+1} = \alpha_t\ (R_{1,t+1} - R_f)\ + R_f \tag{2}$$

其中，$R_{1,t+1} = \exp\ (r_{1,t+1})$ 是单一风险资产的收益，$R_f = \exp\ (r_f)$ 是无风险利率，α_t 是风险资产的份额，无风险资产的份额为 $(1-\alpha_t)$。收入的自然对数服从随机游走，漂移项为 g，假设预期超额（对数）收益 $(r_{1,t+1} - r_f)$ 为常数：

$$Y_{t+1} = Y_t \exp\ (g + \varepsilon_t) \tag{3a}$$

$$r_{1,t+1} - r_f = \mu + u_{t+1} \tag{3b}$$

$$\mathrm{Var}_t\ (u_{t+1})\ = \sigma_u^2 \qquad \mathrm{Var}_t\ (Y_{t+1})\ = \sigma_\varepsilon^2 \qquad \mathrm{Cov}_t\ (u_{t+1},\ \varepsilon_{t+1})\ \sigma_{\varepsilon u} \tag{3c}$$

误差项 u_{t+1} 独立于就业/退休的状态。此外，假设消费增长和资产收益服从联合对数正态分布。

就业/退休状态是随机的。就业的概率为 π^e，退休的概率为 $\pi^r = \ (1-\pi^e)\ \ (0<\pi^r<1)$，退休状态不可逆转，退休时的劳动收入为零。在退休后，死亡的概率为 π^d，因此人们在退休后的寿命为 $1/\pi^d$ 年，退休之前的预期工作年数为 $1/\ (1-\pi^e)$。

因为退休时的收入为零（用上标 r 表示），所以我们有欧拉方程的标准形式

$$1 = E_t\ \{\theta^r\ (C_{t+1}^r / C_t^r)^{-\gamma} R_{i,t+1}\} \tag{4}$$

其中，$\theta^r = \ (1-\pi^r)\ \theta$，$R_{i,t+1}$ 中的 i 为 1、f 或 p，也就是说，$R_{i,t+1}$ 是风险资产、无风险资产或投资组合的收益。Viceira 在二阶 Taylor 级数展开式的基础上给出了近似解，因此他考虑了谨慎储蓄效应（即波动率项）（Campbell 和 Viceira（1999））。这种方法为我们提供了一些有用的直觉，使得消费和资产份额存在闭型解。对退休状态而言，最优（对数）消费和资产份额为

$$c_t^r = b_0^r + b_1^r w_t \tag{5a}$$

$$\alpha_t^r = \frac{\mu + \sigma_u^2/2}{\gamma b_1^r \sigma_u^2} \tag{5b}$$

其中，

$$b_1^r = 1 \tag{5c}$$

$$b_0^r = \ln\left\{1 - \exp\left[\left(\frac{1}{\gamma} - b_1^r\right) E_t r_{p,t+1}^r + \frac{1}{\gamma}\ln\theta^r + \frac{1}{2\gamma}\ (1 - b_1^r\gamma)^2 \mathrm{Var}_t(r_{p,t+1}^r)\right]\right\} \tag{5d}$$

我们可以将（5a）与前面的结论（参见 Samuelson（1969）和第 15 章中的方程（45））联系起来

$$\left(\frac{C}{W}\right)_t = \frac{1}{1 + \ [\theta^r E_t\ (R_{p,t+1}^{1-\gamma})\ ]^{1/\gamma}}$$

可以证明，当 $R_{p,t+1}$ 服从对数正态分布时，上述结论可以简化 Viceira 的解 $c_t - w_t = b_0^r$（参见 Viceira（2001），附录 B）。常数项 b_0 取决于 γ、θ、预期收益和收益的条件方差。但是，如果 $\gamma = 1$（对数效用），那么根据（5d），（对数）消费－财富比率只取决于 θ，与预期收益和收益的方差独立。这反映我们在前面所得到的结论，$(C/W)_t = [(1-\theta)/(1-\theta^{T-t+1})]$。

根据（5d），消费－财富比率是时间偏好率 $-\ln\theta \approx t_p$ 的递增函数，消费－财富比率与投资组合预期收益之间的关系取决于 CRRA 参数 γ 是大于 1 还是小于 1。对偏好为对数效用的投资者而言，$(C/W)_t$ 不依赖于收益。对更多的风险规避类型投资者而言（即 $\gamma > 1$），$(C/W)_t$ 是投资组合预期收益中的递增函数（参见（5a）和（5d）），因为 Er_p 增加所导致的收入效应大于替代效应）。当 $\gamma = 1$ 时，$\mathrm{Var}_t(r_{p,t+1})$ 对 $(C/W)_t$ 的影响为零，但这种影响在其他情形中为负（r_p 的不确定性越大，$(C/W)_t$ 越小，当前储蓄（即谨慎储蓄）就越高）。

因为收益独立同分布（劳动收入为零），所以风险资产的份额 α_t^r 在所有退休时期中都相等（即通常的近视结论）。但 α_t^r 越高，预期收益 μ 就越大，收益的波动率 σ_u^2 就越低。

对冲需求

到目前为止，这些都是在重复前面的结论。但是，在就业状态下消费和资产份额到底怎样呢？在就业状态下一种可能是，继续被雇佣并获得劳动收入，另一种可能是你会进入退休时期。因此，欧拉方程为

$$1 = E_t\{[\pi^e\theta^e(C_{t+1}^e/C_t^e)^{-\gamma} + (1-\pi^e)\theta^e(C_{t+1}^r/C_t^r)^{-\gamma}]R_{i,t+1}\} \tag{6}$$

i 为 1、f 或 p，$\theta^r = (1-\pi^r)\theta$，$\theta^e = \theta$。（对数）消费和风险资产份额的解为

$$c_t^e - y_t = b_0^e + b_1^e(w_t - y_t) \tag{7a}$$

$$\alpha^e = \frac{\mu + \sigma_u^2/2}{\gamma\bar{b}_1\sigma_u^2} - \frac{\pi^e(1-b_1^e)}{\bar{b}_1}\frac{\sigma_{\varepsilon u}}{\sigma_u^2} \tag{7b}$$

其中，$0 < b_1^e < 1$，$\bar{b}_1 = \pi^e b_1^e + (1-\pi^e)$，$b_1^e$ 是关于 g（收入增长率）的递增函数，也取决于收入增长的方差、投资组合收益的方差及两者之间的协方差。α^e 对到退休前的预期工作年数 $(1-\pi^e)^{-1}$ 的依赖性意味着（就业时）资产份额对所考虑的“期限”存在倚赖性。

考虑 $\sigma_{\varepsilon u}$ 的特殊情形，收入冲击与资产收益冲击无关。于是，可以证明 $\alpha^e > \alpha^r$。我们也可以证明如下结论，退休前的预期工作年数越大，α^e 就越大。即使收益独立同分布，这个结论也成立（参见 Viceira（2001），表 1）。表 1 对 $\gamma = 2$ 和 5 时的结果进行了说明。当 $\gamma = 5$ 时，当退休前年数由 5 增加到 35 时，α^e 从 42% 增加到 76%，而 α^r（$\gamma = 2$）为 34%。对给定的退休前年数而言，风险规避程度（γ）的提高会导致 α^e（和 α^r）值的下降。

表 1　当 γ 和退休前的预期时间发生改变时投资组合份额 α^e 的大小

	退休前的预期时间（年）				
CRRA（γ）	35	20	10	5	α^r
	Panel A：Corr（$r_{i,t+1}$，Δy_{t+1}）＝0%				
2	294	190	140	115	90
5	76	61	49	42	36
	Panel A：Corr（$r_{i,t+1}$，Δy_{t+1}）＝25%				
2	255	175	133	111	89
5	63	52	44	39	36

资料来源：摘自于 Viceira 2001 表 1。

因此，最优解与我们的经验规则——工作人员应该比退休人员拥有更多的风险资产——一致。但是，这一经验规则是否稳健呢？如果我们考虑对冲需求，即（7b）右边的第二项，那么答案就是否定的。“对冲项”为$\beta = \sigma_{\varepsilon u}/\sigma_u^2$，它是劳动收入冲击对预料不到股票收益进行回归的斜率。当收入冲击和收益冲击高度负相关时，α^e 会降低，你会持有更少的风险资产，因为在收入较低时资产的收益也较低。因此，资产对消费的匀滑作用很小。将表 1 中的面板 A 和面板 B 进行比较就可以看出，劳动收入与股票收益之间的正相关性是如何降低 α^e 的（对给定的任意退休前年数和 γ 而言）。

反之，如果 $\sigma_{\varepsilon u} < 0$，那么风险资产可以很好地对冲不利的收益冲击所导致的风险，从而对冲需求为正（即 $\alpha^e > \alpha^r$）。在我们前面的简单模型中，劳动收入不是随机的，因此对冲需求不会出现，这里的对冲需求对 α^e 有着非常强大的影响，这种对冲效应越大，退休前的预期年数就越大。然而，当收入冲击和资产收益冲击之间的相关性增加 25% 时（且 γ 介于 2 和 12 之间），$\alpha^e > \alpha^r$（但是，与零相关性的情形相比，$\alpha^e - \alpha^r$ 更小一些）——参见 Viceira（2001）表 1 中的面板 B。这个模型在数值模拟的基础上给出了一些有趣的结论。但我们现在要转为讨论模型的另外一种变体，新的模型将讨论社会保障税收是否可以投资于股票和无风险资产。

社会保险基金和养老基金

美国社会保险基金目前投资于各种无风险债券。然而，如果投资者个体能够借入权益资产，那么他们就可以实现一个投资组合，这个投资组合的风险与投资于权益资产的退休金账户的风险相同。因此，只有当一些（穷）人无法借入权益资产或面临参与股票市场的高昂成本时，将社保基金投资到权益资产上才会增加社会福利。

Campbell 和 Viceira（1999）的思想是对事前的终生福利收益进度量，方法是将 ψ% 的社会基金投资于股票，将（$1-\psi$）% 投资于无风险资产，其中 ψ 在 0 到 100% 中取值。收益独立同分布，收益冲击与收入冲击同期相关。收入过程比 Viceira 中的确定性生命周期效应（如收入关于年龄具有驼峰形状）及持久和暂时冲击的情形更加复杂。为了规避未来的劳动收入风险和退休后的财富风险，投资者不可能进行借贷，也不可能卖空股票。在社保基金中强制储蓄 ψY_{it}后，投资者的可支配收入为

$$Y_{it}^d = (1-\psi)\ Y_{it} \tag{8}$$

退休缴款 ψY_{it}可以投资于股票 R_1 或无风险资产，这与任何个体 i 的流动财富 W_{it}^L一样：

$$W_{i,t+1}^R = R_{p,i,t+1}^R\ (W_{it}^R + \psi Y_{it}) \tag{8a}$$

$$W_{i,t+1}^L = R_{p,i,t+1}^L\ [W_{it}^L + (1-\psi)\ Y_{it} - C_{it}] \tag{8b}$$

$$R_{p,i,t+1}^k = \alpha_{it}^k R_{1,t+1} + (1-\alpha_{it}^k)\ R_f \tag{8c}$$

其中，$k = L$（流动财富）或 R（退休财富）。最大值函数是消费的跨期幂效用函数（用“活着的”条件概率进行加权），个体选择 α_{it}^L，α_{it}^L依赖于状态变量［W_{it}^L +（$1-\psi$）Y_{it}］和 W_{it}^R。我们可以得到 ψ 和 α_{it}^R各种取值下 α_{it}^L和 C_{it}的数值最优解，其中决定社会保障税，α_{it}^R是社保基金投资于股票的比例。固定的“终点”是为了保证，退休群体的平均（个体之间进行平均）收入替代率为 60%。因此，当实际利率为 2% 时，更多的社会保险退休基金投资股票的平均收益率为 6%（即股权溢价为 4%），这意味着降低社会保障税是可行的（固定的替代比率为 60%）。事实上，α^R 从 0 上升到 50% 将使 ψ 从 10% 下降到 6%。

Campbell 等人发现，如果社保基金的一半投资于股票，那么福利收益大致为 3.7% 的消费增加，但是，如果社会保障税为常数，那么福利收益就降为 0.5%。他们也承认这个模型相对于复杂模型的局限性。首先，即使退休后的财富更多地投资于股票，股权溢价也为常数。事实上，降低股权溢价的福利效应取决于股权溢价是由风险资产收益 R_t 的下降所导致的还是由无风险利率 r 的上升所导致的。一个非常明显的结论是，R 的下降和 r 的上升都会减少将社保基金投资从债券转移到股票所到来的福利收益。

模型体现了具有“自融资”性质的社会保障退休系统，因此，家庭之间不存在重新分配。

一些群体将因为社会保障退休基金的终端财富值太小而遭受损失，因为这将导致一个很低的退休年金。这会导致他人推测这部分人退出的可能性，从而影响他们的消费 - 储蓄模式。收益独立同分布，所以不存在择时交易的可能性（参见下面的章节），房产价值已被排除在外。另外，异质型劳动收入冲击与收益不相关，劳动供给是外生的。然而，这个模型在重要政策的评价和思考方面迈出了非常有用的第一步。

16.2 多种风险资产

我们已经看到，引入一个额外的随机变量（即劳动收入）就产生了风险资产的对冲需求。现在，我们允许投资者持有一种以上的风险资产（Campbell 和 Viceira 1999），从而对冲需求存在，因为不同资产收益之间存在条件相关性。我们的模型在 Var 系统基础上允许资产收益具有可预测性。Campbell、Chan 和 Viceira（2003）假设投资者具有 Epstein - Zin 递归偏好、无劳动收入、无投资组合约束（例如，允许卖空，不存在借贷约束）。这种方法的优美之处在于 Campbell 等人给出了近似的解析解，因此我们能获得一些直觉。效用函数为

$$U\left(C_t,\ E_t\left(U_{t+1}\right)\right)=\left[(1-\delta)\ C_t^{\frac{1-\gamma}{\theta}}+\delta\left(E_t\left(U_{t+1}^{1-\gamma}\right)\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}} \tag{9}$$

其中，$\gamma>0$ 是相对风险规避系数，$\psi>0$ 是跨期替代弹性，$0<\delta<1$ 是折现因子，$\theta=(1-\gamma)/(1-\psi)^{-1}$。对时间可分离的幂效用函数而言，$\gamma=\psi^{-1}$，因此 $\theta=1$。如果是对数效用函数，则有 $\gamma=\psi^{-1}=1$。

预算约束和投资组合的实际收益为

$$W_{t+1}=\left(W_t-C_t\right)R_{p,t+1} \tag{10a}$$

$$R_{p,t+1}=\sum_{i=1}^{k-1}\alpha_{i,t}\left(R_{i,t+1}-R_{k,t+1}\right)+R_{k,t+1} \tag{10b}$$

其中，第 k 种资产为基准资产（如短期国债），但它不是实际的无风险利率（既存在通货膨胀风险）。对投资组合收益和预算约束进行对数线性处理之后，资产 $i=1,2,\cdots,k-1$ 相应的欧拉方程为

$$\begin{aligned}&E_t\left(r_{i,t+1}-r_{k,t+1}\right)+\frac{1}{2}\operatorname{Var}_t\left(r_{i,t+1}-r_{k,t+1}\right)\\&=\frac{\theta}{\psi}\left(\sigma_{i,c-w,t}-\sigma_{k,c-w,t}\right)+\gamma\left(\sigma_{i,p,t}-\sigma_{k,p,t}\right)-\left(\sigma_{i,k,t}-\sigma_{k,k,t}\right)\end{aligned} \tag{11}$$

小写字母表示取对数后的值。因此，资产 i 的预期超额收益依赖于资产 i 的收益与（ⅰ）（对数）消费 - 财富比率 $c-w$、（ⅱ）投资组合的收益和（ⅲ）基准资产 k 收益之间的条件协方差。（11）中的方差项为 Jensen 不等式项，如果消费和资产收益服从联合对数正态分布，即 $\psi=1$ 的情形，那么（11）就是精确的等式。可以证明，当资产收益（和任意强制（forcing）变量，如红利价格比）由带有同方差误差的线性 Var 系统所驱动时，协方差就是 Var 中状态变量 z_t 的线性函数。同方差意味着状态变量只能预测预期收益的变化，无法预测风险的变化（但是，风险与平均收益之间的联系在实证研究中不是很强——Chacko 和 Viceira（1999））。求解欧拉方程就得到最优风险资产份额：

$$\begin{aligned}\alpha_t=&\frac{1}{\gamma}\sum\nolimits_{xx}^{-1}\left[E_t\left(x_{t+1}\right)+\frac{1}{2}\operatorname{Var}_t\left(x_{t+1}\right)+(1-\gamma)\ \sigma_{k,x}\right]\\&+\frac{1}{\gamma}\sum\nolimits_{xx}^{-1}\left[-\left(\frac{\theta}{\psi}\right)\left(\sigma_{i,c-w,t}-\sigma_{k,c-w,t}\right)\right]\end{aligned} \tag{12}$$

其中，x_{t+1} 是超额收益 r_i-r_k（$i=1,2,\cdots,k-1$）构成的向量，协方差矩阵为 Σ_{xx}。方括号中的项是 Var 中变量 z_t 的线性函数，$z_t=\left[r_{k,t},\ x_t,\ s_t\right]$ 和 s_t 是 Var 中的“非收益性”变量（如红利价格比）。因此，当 Var 中的变量存在可预测性时，资产份额具有时变性。

（12）右边第一项是资产需求的近视部分（myopic component），因为它只依赖于下期的

（超额）收益 x_{t+1}，不依赖于跨期替代弹性 ψ。对具有对数效用偏好的投资者而言，$\gamma=1$，$\theta=0$，于是括号中第二项（对冲需求）为0，剩下的只是近视需求（myopic demand）。如果投资机会集合为一个常数，那么随着时间的推移，对任意 γ 值而言，对冲需求都为0。另外，如果基准资产是无风险资产，那么 $\sigma_{k,x}=0$，资产份额熟悉的静态解为 $\alpha_t=\gamma^{-1}\sum_{xx}^{-1}[E_t x_{t+1}+(1/2)\mathrm{Var}_t(x_{t+1})]$。一般而言，资产收益会发生变化，因为 $E_t(x_{t+1})$ 和协方差项（线性地依赖于 z_t）的预测值都会发生变化。消费增长方程为

$$E_t\Delta c_{t+1}=\psi\ln(\delta)+\frac{1}{2}\frac{\theta}{\psi}\mathrm{Var}_t(\Delta c_{t+1}-\psi r_{p,t+1})+\psi E_t(r_{p,t+1}) \tag{13}$$

在（13）中，消费增长越高，δ 就越高（即当前消费的权重越低），投资组合的预期收益越高，$\mathrm{Var}_t(\Delta c_{t+1}-\psi r_{p,t+1})$ 越低（因为 $\theta>0$）——这对幂效用函数而言也成立，因为 $\theta=1$。

如果偏好为递归偏好，那么只是在一些特殊情形（Giovannini 和 Weil（1989）），我们才能得到闭型解。我们将结论与前面的结论进行如下比较：

* 对 *iid* 收益（Var 只包含一个常数）而言，投资者将总财富的一个固定比例用于消费，具体比例取决于他的时间偏好率、相对风险规避程度、跨期替代弹性 ψ 和投资年限 $T-t$（Bhamra 和 Uppal（2002））。与此同时，也存在一个近视的投资组合规则，也就是说，投资者选择 α_t 时就像把投资期限视为单期一样（α_t 独立于 ψ）。

* 如果预期收益具有时变性但 $\gamma=1$，那么近视的投资组合规则也适用，但消费－财富比率具有时变性。如果 $\gamma=1$ 且 $\psi=1$，那么消费－财富比率是一个常数。

* 如果 $\gamma=1$ 但 $\psi\neq1$，那么消费－财富比率为常数（$=1-\delta$），但最优投资组合规则不是近视的。

Campbell、Chan 和 Viceira（2003）使用数值方法对模型进行求解，三种风险资产是股票、（名义）债券和短期国库券。如果 Var 只包含一个常数项（即收益独立同分布）且 $\gamma=1$，那么最优投资组合份额是近视的，只依赖于方差－协方差矩阵 Σ_{xx} 和各种资产的（常数）预期超额收益。在这种情形中，投资者只有股票和债券的多方头寸，股票的持有比例更高一些，因为无条件预期收益更高一些（即事后的股权溢价很高）。以此为基准，他们将好几个变量纳入到 Var 模型中并考虑 $\gamma\neq1$ 的情形。对冲需求在这里会影响 α_t 的均值和 α_t 对状态变量 z_t 的敏感性。对风险规避系数 $\gamma>1$（即风险容忍能力 $=1/\gamma<1$）的保守型投资者而言，股票需求比近视投资组合需求大，因此对冲需求为正（且是 $1/\gamma$ 的驼峰形函数）。事实上，股票对未来收益来说是一种很好的对冲工具，因为（Var 模型中）红利价格比冲击与股票收益冲击负相关，这会导致股票对冲需求的增加（当 $\gamma>1$ 时）。他们也发现，（名义）债券冲击与股票收益冲击负相关，这会鼓励（风险规避程度为中等）投资者做空债券，目的是为了持有股票的多方头寸。当我们将通货膨胀关联债券纳入到资产表中时，非常保守的投资者会持有这种资产的多方头寸，目的是为了对冲实际利率风险，因为这种资产是股票头寸的良好对冲工具（即指数债券冲击和股票收益冲击之间具有负相关性）。表2对主要的跨期消费－投资组合模型进行了概括。

16.3　不同偏好

我们在前面的章节中已经注意到，损失规避偏好（loss aversion（LA）preference）——加上幂效用函数（多期消费——Barberis 和 Huang（2001））——有助于用合理的风险规避参数值来解释股权溢价之谜。引入损失规避的原因之一是，取决于消费水平的相对风险规避（CRRA）效用函数（如幂效用函数）存在很多问题。

第一，它所给出的股权溢价很低。第二，如果股票的预期收益为正，那么所有投资者都将持有股票，但是，在绝大多数国家中，只有很少一部分人在股票上进行投资。对 CRRA 投资者

而言，降低股票市场参与度的唯一方法是假设投资者有很高的（固定）进入成本（Campbell 和 Viceira（1999），Liu 和 Loewenstein（2002））。第三，CRRA 和期望效用框架与“Rabin 之谜”（Rabin（2000））不一致。如果（追求期望效用最大化的）投资者拒绝以相同概率输 100 美元或赢 110 美元的“小额”赌博（对任意初始财富水平而言），那么他们也会拒绝以 50:50 概率损失 100 美元或赢任意数量美元的赌博，这就是“Rabin 赌博”（Rabin（2000））。例如，在 CRRA 偏好和 $\gamma=10$ 情形中，投资者会拒绝 111/110 赌博。然而，为了规避一个损失 1000 美元或得到 1 百万多美元（同等概率）的赌博，她的支付意愿（即不参与赌博的确定性财富减去参与赌博后的确定性等价财富）仅仅是 280 美元左右。随着财富的增加，CRRA 效用函数的边际（财富）效用趋于递减，并很快地降为零。因此，CRRA 投资者会以一个非常小的代价来拒绝参与 Rabin 赌博。

与标准期望效用型投资者不同，为了规避 100/110 赌博，损失规避型投资者的效用函数不必具有很大的凹性，因为效用函数在参考点上不连续。损失规避意味着他们痛恨失去 100 美元，但对收益而言效用函数不需要很大的曲率。尽管如此，正如我们已经看到的那样，损失规避有助于解释股权溢价之谜，Ang 和 Bekaert（2001）指出了损失规避偏好的一些缺陷，包括“参考点”选择的任意性（由此可以度量收益和损失）和如何更新“参考点”。在应用投资组合理论的时，对风险资产的份额而言可能得到一些“奇怪”的结论（例如，可能不存在有限的最优权重，对零风险溢价而言 α^* 可能小于零）。

表 2　　　　资产需求：跨期消费投资组合选择

		劳动收入	收益 IID	可预测收益	消费
Samuelson（1969）	重新进行调整离散时间幂效用	无	$\alpha^*=(ER-r)/\gamma\sigma^2$ α^* 与期限独立（近视）	否	$C/W=f(ER,\gamma,T-t)$ $C/W=f(\gamma,T-t)$——对数效用
Merton（1969，1971）	重新进行调整连续时间幂效用	无	与 Samuelson（1969）相同	加上对冲需求——预期收益与收益冲击之间的协方差	与 Samuelson（1969）相同
Campell 和 Viceira（1999）	重新进行调整离散时间	无	对任意风险资产而言，$r=1$ 时 $\alpha^*=f(\sum^{-1},ER,\gamma,\psi)$	加上对冲需求——不同资产	与 Samuelson（1969）相同
Bhamra 和 Uppal（2002）	递归效用对数线性近似		$\gamma=1$ 时是近视的	收益冲击与 C/W 比之间的协方差	但 $\psi=1$ 时 C/W 为常数
Brodie、Merton 和 Samuelson（1992）	与 Samuelson（1969）相同	定性	近视，但 α^*（有劳动收入）> α^*（无劳动收入），劳动收入与无风险资产相似，对更多的风险资产也成立	否	与 Samuelson（1969）相同
Viceira（2001）	重新进行调整离散时间幂效用对数线性	随机劳动收入	近似需求加上对冲需求（收入冲击与劳动收入冲击之间的协方差）	否	与 Samuelson（1969）相同，加上 Var（R）中的 Jensen 不等式项。对对数效用函数而言 C/W 与 ER 独立

注：此表对已有研究的主要结果进行了概括，具体细节请参见正文。

Ang 和 Bekaert（2001）建议使用失望规避（DA，disappointment - aversion）偏好（Gul（1991）），这意味着投资者希望规避损失，但参考点是内生的，参考点定义为确定性等价之上或之下的结果。CRRA 效用函数是 DA 偏好的特殊情形。因为 DA 偏好包含损失规避，他们就避免了“Rabin 之谜”。例如，对 $\gamma=2$ 的 DA 投资者而言，1.18 的效用损失比同等收益对他的影响更大，他会拒绝一个 100/110 的赌博但愿意参加大额 Rabin 赌博（即愿意支付 3664 美元来规避 50:50 赢 25000 美元或损失 2000 美元的赌博）。

在投资者追求期末（财富）效用最大化且具有 DA 偏好的模型中，我们能解释为什么股票市场的参与度很低。与收益相比，人们不喜欢同样大的损失，很多人会选择根本不持有股票（如果他们的合理损失规避程度为 1.66 左右）。与此相对照的是，CRRA 效用永远也不会给出不投资股票的配置策略（即 $ER-r>0$）。当 $\alpha^*=5\%$ 时，γ 高达 36。如果多个国家的熊市正相关，那么失望规避和本土偏好（home bias）一致，因为 DA 投资者不喜欢负的偏态，倾向于减少跨国持有的股票。DA 投资者也希望预防衰退，因此他们愿意购买看跌保护型产品。DA 效用在投资组合配置实践中的应用相对较新，它是否能解释较大范围内的一些现象，我们还需拭目以待。

16.4　期限效应与不确定性

到目前为止，就投资期限如何影响投资者对股票和债券（无风险资产）的选择而言，我们只谈了一点点。在静态均值 - 方差模型中，风险资产的最优投资比例是条件期望收益（相对于投资组合收益的方差而言）来决定，但投资期限由投资者外生选择。这就是单期最大化问题。在每一期中，投资者都要进行新的预测，确定“新的”投资比例。在收益 *iid* 的特殊情形中（因此，投资组合的收益及收益的协方差都是常数），风险资产的投资比例不会随着时期的演变而发生改变。因此，最优资产比例对投资期限不同的两个投资者而言是相同的。从直觉来看，这是因为独立同分布的（对数）资产收益意味着预期收益和投资组合的方差都是投资期限的递增函数（即 $ER_T=T\mu$，$\sigma_T^2=T\sigma^2$）。在任何情形中，均值 - 方差模型都不是跨期模型，因此它对投资期限效应的连贯考察而言不是特别有用。

Barberis 和 Huang（2001）在幂效用（取 $\gamma=10$）函数假设下考察期限效应，模型中的代表性投资者在时刻选择资产比例来使期末（财富）效用 $EU(W_{T+H})$ 到达最大，所以投资组合在 T 到 $T+H$ 之间是不会进行重新调整的（模型中也没劳动收入）。

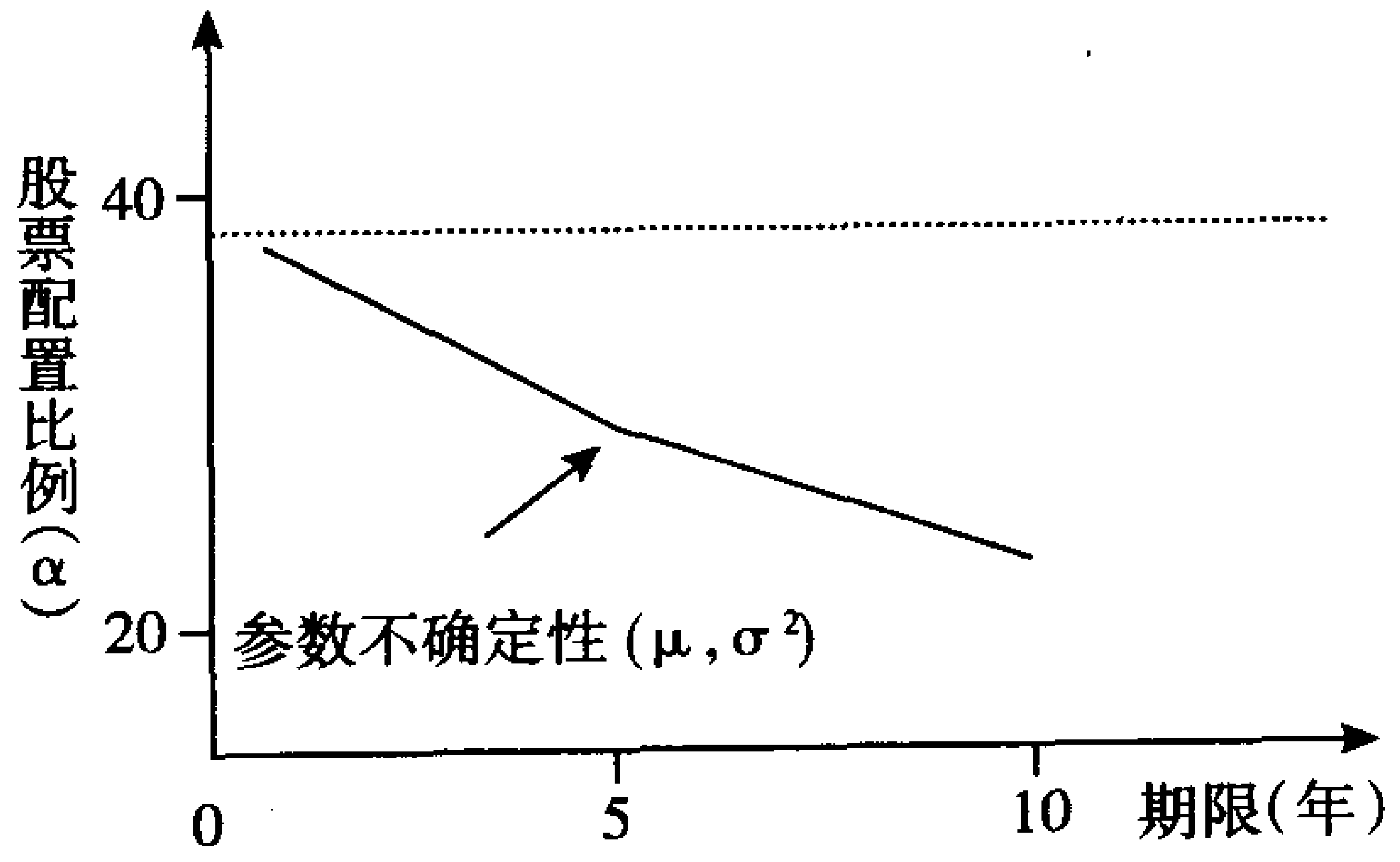


图 1　**“买而持有”：无可预测性（$\gamma=10$）**

注：1. 收益模型为 $r_{t+1}=\mu+\varepsilon_{t+1}$，$\varepsilon_{t+1}\sim N(0,\sigma^2)$。

2. 对（μ，σ^2）进行估计的数据时期为 1986—1995。

买而持有，收益独立同分布

当效用函数为幂效用函数、收益独立同分布且投资组合进行连续调整时，最优资产份额与期限 H 独立（Samuelson（1969））。部分原因是在于如下：独立同分布收益、长期平均收益 $H\mu$ 和方差 $H\sigma^2$ 都是期限的线性函数。如果遵循买而持有策略且参数没有不确定性，那么结论继续成立。当 $\gamma=10$ 时，我们发现 $\alpha^* \approx 40\%$（参见图1）。如果我们现在引入参数不确定性，α^* 与期限 H 负相关且 α^* 依赖于 H，那么出现的结果是 α^* 从 α^*（$H=1$）$\approx 35\%$ 单调下降到 α^*（$H=10$）$\approx 20\%$（Barberis 和 Huang（2001），图1）。这是因为，如果 $R_T=\mu+\varepsilon_t$，μ 不确定但投资者可以用贝叶斯规则来对它的估计值进行更新，那么这就在长期限中引入了其他风险，从投资者角度来看收益将不再独立同分布。例如，如果存在一些好（坏）的收益，那么贝叶斯更新会增加（降低）代表性投资者对 μ 的估计，收益中也就引入了正自相关性。后者意味着 $\sigma_T^2 > T\sigma^2$，股票的持有比例 α 随着投资期限的增加而下降。

买而持有：收益可预测

股票收益和红利价格比的预测方程为

$$r_{t+1}=\mu+\beta z_t+\varepsilon_{r,t+1}$$

$$z_{t+1}=\delta+\phi z_t+\varepsilon_{z,t+1}$$

我们所要求的分布取决于 T 时刻红利收益 z_T 的值，T 时刻是买而持有的开始时刻。现在，多期收益的方差取决于收益预测方程和红利收益方程中误差项之间的同期协方差，因为（例如）

$$\mathrm{Var}_T\ (r_{T+1}+r_{T+2})\ =2\sigma_r^2+\ (\beta^2\sigma_z^2+2\beta\sigma_{r,z})$$

基于美国数据的实证结果表明，$(\beta,\ \phi)\ >0$ 且 $\sigma_{r,z}<0$，括号中的项为负。因此，可预测性但无参数不确定性意味着长期限条件方差是投资期限的递增函数，但增长速度缓慢（即 $\sigma_H^2 < H\sigma^2$）。

记住我们在前面所看到的结论，在无参数不确定性和收益不可预测（即 *iid* 收益）情形中，α^* 与投资期限独立。然而，在无参数不确定性但收益可预时（$\gamma=10$），最优 α^* 随着投资期限 H 的增加而快速上升，从 $H=1$ 时的 $\alpha^* \approx 40\%$ 上升到 $H=10$ 时 $\alpha^* \approx 90\%$（参见图2）。原因有两个方面。首先，时刻的红利价格比能预测未来的高收益。其次，因为 $\sigma_H^2 < H\sigma^2$，所以股票对长期投资者而言风险更低。这两个方面的原因导致长期限投资者增加股票持有量。

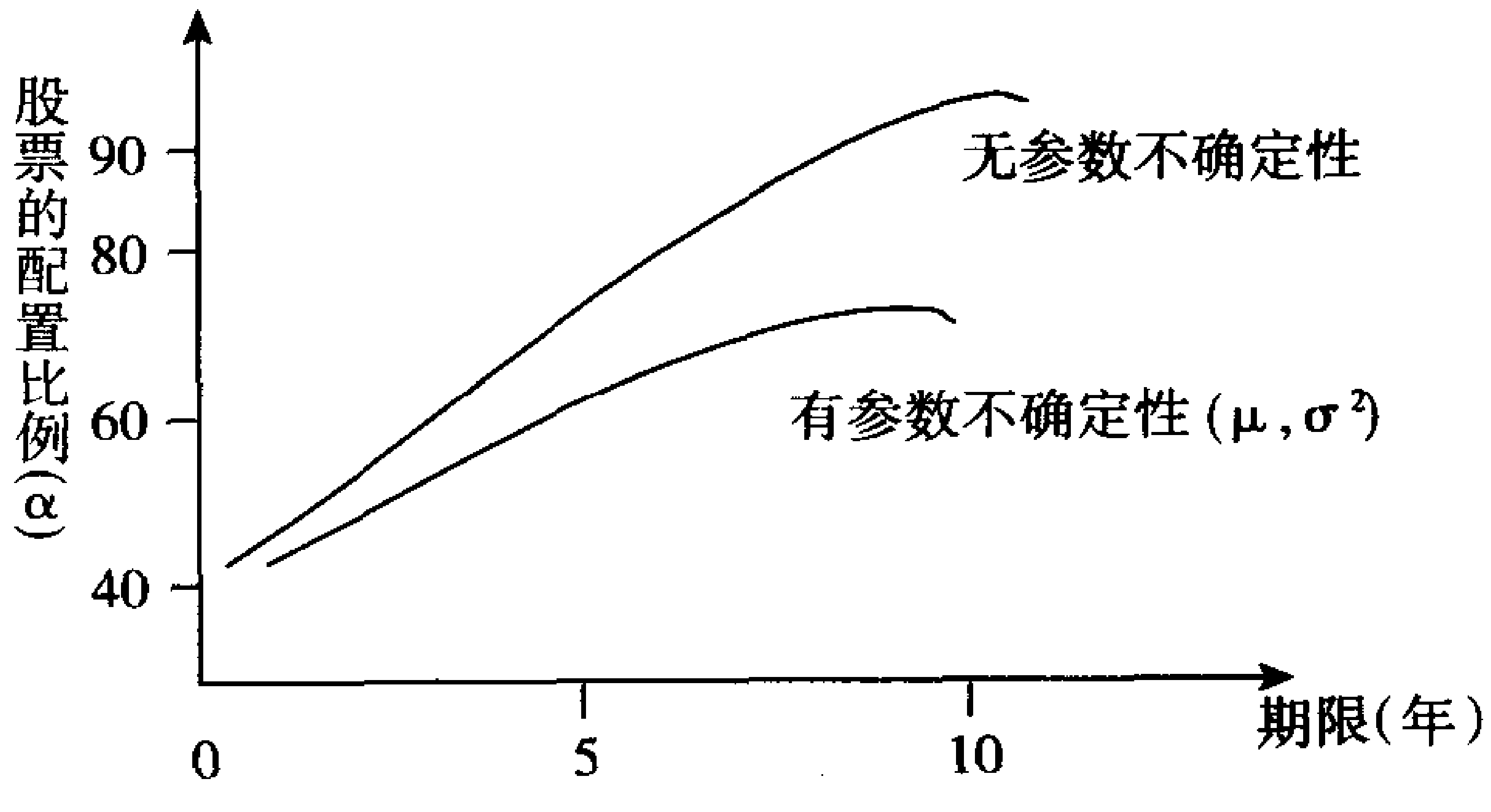


图2　“买而持有”：收益可预测（$\gamma=10$）

注：1. 对收益和红利价格比使用双变量 Var 模型（1952—1995）。

2. $T=0$ 时刻的红利收益设定为样本均值。

然而，如果我们现在考虑参数不确定性，那么最优 α^*（见附录）仍然随着投资期限的增加而上升（一直到第7年），然后略有下降，渐近靠近一个更低的值 α^* =（$H=10$）$\approx 55\%$，而“无不确定性”情形中的值为 α^*（$H=10$）$\approx 90\%$（图2）。就长期而言，最优值 α^*（$H=$

10）出现下降的原因有两个方面。在无不确定性时，σ_H^2 的增长速度比（期限）线性增长速度慢，这倾向于增加股票需求。但是，如果参数存在不确定性和效用函数存在凹性（即风险规避），那么收益不确定性的“下降”效应比（同样大小的）上升效应更大一些。因此，投资者不确定红利收益是否“减缓”了条件方差（关于期限）的增长速度（与无不确定性的情形相比，这使股票的风险看上去更大一些，也使 α^* 的值有所减小）。表 3 对这些结果进行了概括。

表 3　　期限效应与不确定性：买而持有

		风险资产的需求	原因
	Panel1：IID 收益		
1A	无参数不确定性	α 与期限独立	$\mu_H = H\mu$，$\sigma_H^2 = H\sigma^2$ 预期收益和方差都是期限的线性递增函数
1B	参数不确定性，(μ，σ^2)	α 是期限的递减函数	σ_H^2 比期限本身的增长速度快，因为存在估计风险
	Panel2：可预测性（$r_{t+1} = \mu + \beta z_{t+1} + \varepsilon_{t+1}$，且 $z_{t+1} = \delta + \phi z_t + v_{t+1}$）		
2A	无参数不确定性	α 是期限的递增函数	(ⅰ) 大 z_t 表示高未来收益； (ⅱ) $\sigma_{r,z} < 0$，因此 $\sigma_H^2 < H\sigma^2$
2B	参数不确定性	α 是期限的递增函数，但速度没有“无不确定性”情形快	(ⅰ) 关于 μ 的不确定性使得 $\sigma_H^2 > H\sigma^2$，这倾向于减小 α；(ⅱ) z_t 的实际预测功效不确定，这在长期倾向于减小 α（与无不确定性情形相比）

注：投资者在 $t=0$ 时刻选择 α 来使 $U(W_H)$（幂效用函数，$\gamma=10$）达到最大。

动态资产配置：每年重新进行调整

当收益可预测、均值反转现象存在、无参数不确定性时，投资者（$\gamma > 1$）会对投资组合进行重新调整，将更多的资源配置到长期限的风险资产上，以此来对冲预期收益的变化（Merton（1973））。Barberis 和 Huang（2001）认为，当收益可预测但参数存在不确定性时（在进行重新调整的那些时期中不存在学习效应），资产份额 α^* 通常是投资期限的递增函数，但与确定性情形相比，增长速度要慢一些。

Barberis 的模型假设投资者知道收益和红利的函数形式，潜在的真实参数为常数。这些条件在实践中不可能成立，增加的不确定性可能会进一步降低投资者（他考虑参数不确定性）在长期限中的股票最优配置比例。

从整体来看，与无不确定性的情形相比，估计风险会大大地改变最优资产配置（因此，前者所得出的结论不可能很好地解释投资组合的行为，我们在解释时要特别“谨慎”）。

16.5 择时交易与不确定性

在前面的章节中，我们回答了“对投资期限不同的两个投资者而言，他们在风险资产上的持有比例为多大”（两个投资者要么都考虑要么都不考虑参数不确定性）这一问题。与此相对照的是，择时交易在每个时点上根据时刻所观察到的信号（如红利价格比）来及时地调整投资组合。

Campbell 和 Viceira（1999）在连续调整的跨期（消费）效用函数模型中（没有劳动收入）直接强调了择时交易问题。收益和红利价格比的 Var 模型证实了“可预测性”。Campbell - Viceira 模型有着非常明显的择时交易效应，股票配置比例从 -50% 变动到 220%，红利价格比与均值的偏离程度从 -2 倍标准差变动到 +1 倍标准差。明显的择时交易效应也是其他预测计划（Brennan、Schwartz 和 Lagnado（1997））和方法（Gallant、Hansen 和 Tauchen（1990），Brandt（1998））的共同特征。

显然，这些择时交易结论基本上忽视了预期收益预测的不确定性。Barberis 和 Huang（2001）也对这个问题进行了考察，在他们的模型中，代表性投资者追求（而不是跨期消费）的最大化，中介交易不存在（即买而持有策略）。

在无参数不确定性时，初始红利收益越高（当前价格越低），股票的配置比例（对任何期限而言）就越高，α^* 仍然是期限的递增函数。然而，在存在参数不确定性时（且 $\gamma=10$），T 时刻的高红利收益很难对长期 α^*（$H=10$）产生影响（但没有短期效应）。这种不敏感性的原因是，红利收益的真实预测功能不确定，“信号”在长期不具有可靠性。因此，在无参数不确定性时，α^*（$H=10$）对红利收益的初始值非常敏感，但在投资者考虑参数不确定性时却变得很不敏感。（另外，在无参数不确定性时，如果初始红利价格比非常高，那么风险资产的份额与期限负相关，长期限投资者比短期限投资者持有更少的股票。）

换句话说，当红利价格比偏离均值的程度给定时，参数不确定性的考虑极大地降低了股票的最优配置比例。如果“存在参数不确定性”，对不同的期限而言，最优配置比例（$\gamma=10$）落在 30% 到 50% 的狭窄区域内。与此相对照的是，“无参数不确定性”情形中的最优配置比例落在 30% 到 100%（最大）的区域内。

这些模型中的择时交易也意味着，你会经常错过市场的上升期（如 1995—2000），而被动型指数追踪基金则会把握住这种机会。因此，这里存在另外一种风险（择时交易模型没有对此进行考虑）：在你的择时交易策略产生赢利之前，你可能已经失业。

因此，尽管可预测性意味着在投资组合配置方面存在很明显的择时交易效应，在问题中参数不确定性的考虑会极大地降低这种效应。

16.6 随机参数

如果收益独立同分布且投资者具有幂效用函数，那么跨期（可加）效用（关于消费的函数）最优化所得到的资产需求 α_t 与投资者的投资期限独立（Samuelson（1969））。如果未来收益具有可预测性（如用红利价格比来进行预测），从而投资机会具有时变性，那么跨期对冲需求存在（Merton（1971））。正如我们已经看到的那样，Barberis 和 Huang（2000）使用了期末财富效用，在预测方程中引入了参数不确定性，但未知参数本身不存在时变性（例如，投资者不知道平均收益，但平均收益的真实值永远不会发生变化）。Xia（2001）使用了连续时间跨期消费模型，允许预测方程中的参数 δ_t 和 β_t（如 $R_{t+1}=\delta_t+\beta_t z_t$）具有时变性（两个参数的动态变化都是均值反转扩散过程），计算了最优资产份额（包含对冲需求）。Xia（2001）指出，对冲需求认知在随机参数情形中非常复杂。资产份额依赖于通常的近视（myopic）项（$ER-r$）$/\gamma\sigma_R^2$ 和对冲需求，对冲需求由三个部分（可加性）组成。对冲需求的第一个部分依赖于（$z_t-\bar{z}$），z_t 是预测变量（当未知参数 β 无不确定性时这一项会消失）。第二个对冲需求组成部分取决于 $\sigma_{z,R}$，如果预测变量 z 和风险资产收益不相关，那么这一项为零。最后，第三个部分取决于 $\sigma_{\beta,R}$，这是预防（未知）参数 β 发生随机变化的对冲需求。

可以证明，依赖于 z_t 当前值的 $\sigma_{\beta,R}$ 可能是投资期限的递增或递减函数。因此，传统建议（年轻投资者应持有更多的股票）一般不成立（Xia（2001））。

正如我们已经看到的那样，在参数无不确定性时，在长期（$T-t$）中追求跨期效用最大

化的投资者具有很强的择时交易能力，他们根据预测变量 z_t 当前值的单调反应来进行股市操作。在进行学习之后，这种反应会逐渐减弱，不再是变量 z_t 的单调函数。这是因为，在进行学习之后，由参数不确定性所导致的对冲需求为负，这最终会对 $\partial\alpha_t/\partial z_t$ 产生主导作用。

因此，在随机参数环境中，对冲效应取决于未知参数的当前估计值、不确定性的估计程度、预测变量 z_t 的当前值和投资期限（$T-t$）。Xia（2001）的模型用一个例子来说明这一领域的结果对学习模型的敏感性。投资者在这里甚至知道预测方程 $R_{t+1}=f(z_t)$ 的函数形式和 z_t 的动态过程（考虑均值反转和连续时间理论中的 Ornstein - Uhlenbeck 过程）。投资者也认为预测方程没有区制转换。因此，对“合理的”学习过程进行建模并将跨期消费 - 投资组合选择模型嵌入进来还有很长的路要走。

16.7　稳健性

在上一节中，我们集中精力讨论了参数不确定性和估计不确定性。不确定性的另一种形式是：投资者在使用一个特殊的资产收益模型（如均值反转过程）时却意识到另外一个模型也可以解释收益。稳健性是指投资者选择资产配置策略来避免使用结果最为糟糕的其他模型。他们希望所选择的模型能产生一个很好的结果。假设替代模型也符合要求，如果替代模型也是对现实的正确描述，那它也会给出一个合理的结果。投资者担心最坏的结果（不同模型之间进行比较）出现，使用期望效用最大化或最小化可以对其进行解释（你使不合意结果出现的可能性达到最小（Roy（1952））。因此，投资者会因为模型不确定性而遭受损失，但投资者的 CRRA 参数仍然很低。

Maenhout（2001）在投资者选择消费和资产份额来最大化跨期效用（幂函数）的模型中（不存在劳动收入）强调了稳健性问题。然而，投资者也考虑了其他股票收益模型（这些模型很“相近”或“相似”），这些模型用参数 θ 来度量稳健偏好的强度（$\theta=0$ 就是标准期望效用最大化问题）。对幂效用函数和独立同分布收益而言，“稳健型投资者”持有的最优资产份额为：

$$\alpha^{*}=\frac{1}{(\gamma+\theta)}\left(\frac{ER-r}{\sigma^{2}}\right)$$

因此，合意的稳健性可以降低风险资产的需求（与期望效用情形相比）。消费财富比率也依赖于 θ、$(ER-u)/\sigma^2$ 和 γ。如果收益可预测（如均值反转），那么稳健性会导致正的风险资产对冲需求（即使对 $\gamma=1$ 的对数效用类型投资者而言，结论也成立），对冲需求通常只在 $\gamma\neq1$ 的预期效用框架下才会长生。注意，“稳健型投资者”不会对周边环境进行学习，他们只是为了规避最为糟糕的结果。稳健性思想（即使存在一些缺陷）在选择各种宏观经济模型来追求结果的最大化方面显得非常重要，在投资组合选择领域非常有用。

结论

上述分析指出了跨期模型的复杂性和最优资产份额对收益可预测性及对冲需求（即不同状态变量冲击之间的协方差）的敏感性。如果结论对特殊假设非常敏感，那么模型的实践价值就值得怀疑。换句话说，所有模型都是对现实的典型描述，因此相对平凡的假设有时对主要结果而言至关重要。然而，只有对各种模型进行仔细考察才能对模型的局限性有所了解，最近的研究为这些问题最提供了大量信息。也许，这一复杂而又富有挑战性的领域的主要特征是你必须对各种可能性进行考虑（我们在这一章里对绝大部分可能性进行了讨论）。我们将其总结如下：

* 在离散时间和连续时间中使用有限或无限期界模型。

* 为了得到富有解释力的闭型解，我们需要一些简化技巧（如 Merton（1971），Kim 和 Omberg（1996），Campbell 和 Viceira（1999，2000））或更为复杂的模型。模型需要用数值方

法来求近似解（如果偏离初始条件太远，那么这些方法可能无效）。

* 对效用函数（如幂效用函数与 Epstein - Zin 递归效用函数或损失规避偏好）及其争论（如消费、损失和收益或期末财富，如 Brennan、Schwartz 和 Lagnado（1997））的选择。

* 所施加的约束（如不允许卖空、借贷限制、破产等等，Brennan、Schwartz 和 Lagnado（1997）），我们通常能得到“复杂的”数值解。

* 对状态变量数目的选择（如多少种资产、是否包含随机劳动收入、考虑随机参数的不确定性）。

* 是否允许资产收益具有可预测性，是否假设代表性投资者知道预测方程中的参数，投资者是否对这些时变参数进行学习和了解。采用那种形式来对参数进行更新（纯粹以数据方案为基础（如递归 OLS 和 Kalman 滤波）、区制转换还是贝叶斯学习机制）。

16.8 小结

* 状态变量（如随机劳动收入但不能是收益）的引入使跨期消费 - 投资组合模型的求解变得非常复杂，也导致了对冲需求。如果收益冲击和劳动收入负（正）相关，那么这将导致风险资产需求的增加（减少）。如果是负相关关系，那么风险资产作为一种对冲工具可以预防劳动收入预料不到的下降，使得消费路径更加匀滑，因此，风险资产的需求更高。当收益（投资机会集）具有时变性时，对冲需求也将上升。

* 损失规避（如失望规避偏好）能解释为什么很多投资者根本不参与股票市场（他们害怕损失，从而他们的理性行为是回避这些风险赌博），当市场下跌时，这些风险资产将导致很高的效用损失。如果参与股票市场的进入成本很高，那么这些结果只能在标准幂效用情形中得到。损失规避效用可以避免 Rabin 矛盾，如果投资者避免一个小的（50∶50 的概率）赌博，那么他们也不会参与损失相对较小而收益却无限大的赌博。

* 参数不确定性将极大地降低各个期限中风险资产的需求（相对于确定性情形而言）。在“模型不确定性”存在时也可能出现这种情形，稳健的最大最小准则在资产配置决策中广泛使用。当收益可预测时，参数不确定性意味着资产份额不会以很快的速度上升，但随着投资者期限的增加资产份额会增加到一个很高的水平（这与资产可预测但无参数不确定性的情形相同）。

* 如果股票收益可预测，那么预测变量（如红利价格比）的值对股票需求有很强的正（择时交易）效应。然而，当参数存在不确定性时这种效应会大大减弱。

附录：参数不确定性与贝叶斯定理

Barberis 和 Huang（2001）使用贝叶斯规则对存在估计不确定性时跨期模型中最优资产份额的确定问题进行了非常清晰的描述。这种方法的好几个方面都广为使用，因此我们来勾画这种方法的主要特征。为了确定今天的最优资产配置（如在一种风险资产（股票）和一种无风险资产之间进行配置），我们不得不预测一期或多期资产收益。假定 T 时刻的初始财富为 W_T，我们希望通过投资来使 H 期后的期末财富效用达到最大。如果 $W_T=1$ 美元，α 是股票的配置比例，那么

$$W_{T+H}=(1-\alpha)\exp(r_fH)+\alpha\exp(r_fH+r_{T+1}+r_{T+2}+\cdots+r_{T+H})$$

其中，r_f 等于实际无风险利率，r_T 等于（每期）实际股票收益（两者都用连续复利进行计算）。当效用函数为 CRRA 效用函数时，

$$U(W_{T+H})=\frac{(W_{T+H})^{1-\gamma}}{1-\gamma} \tag{A1}$$

最简单的情形是买而持有策略，投资者选择 α（在 T 时刻进行选择，持有组合直到 $T+H$ 时刻）来使 $E_T U\ (W_{T+H})$ 最大化，

$$\max_{\alpha} E_T \frac{[\ (1-\alpha)\ \exp\ (r_f H)\ +\alpha \exp\ (r_f H + R_{T+H})\ ]^{1-\gamma}}{1-\gamma} \tag{A2}$$

其中，$R_{T+H} = r_{T+1} + r_{T+2} + \cdots + r_{T+H}$。假设股票收益的生成方程为

$$r_{t+1} = \mu + \beta z_t + \varepsilon_{t+1} \tag{A3}$$

其中，z_t = 红利价格比（比如说 AR（1）），ε_t 为 iid $(0,\ \sigma^2)$，$\theta = (\mu,\ \beta)$。如果我们假设投资者相信 OLS 系数 θ 的估计没有误差（从而 $\hat{\theta} = \theta$），那么投资者的问题就变成

$$\max_{\alpha} \int U\ (W_{T+H})\ p\ (R_{T+H} \mid z,\ \hat{\theta})\ \mathrm{d}R_{T+H} \tag{A4}$$

其中，$p\ (R_{T+H} \mid z,\ \hat{\theta})$ 是概率密度函数，假定 $z = (z_1,\ z_2,\ \cdots,\ z_T)$ 是已知的，它是 z 直到今天（$=T$）的观察值，假设 θ 可由 $\hat{\theta}$ 准确地给出。但（A4）没有考虑 θ 估计的不确定性，因此，这会导致 R_{T+H} 的未来估计存在不确定性。考虑 θ 不确定性的一种有用方法是使用贝叶斯后验分布 $p\ (\theta \mid z)$ 来概括 θ 的不确定性，假定已经观察到的数据是给定的。预测分布是只基于所观察到的数据而不是 θ 任意固定值的概率分布

$$p\ (R_{T+H} \mid z) = \int p\ (R_{T+H} \mid \theta,\ z)\ p\ (\theta \mid z)\ \mathrm{d}\theta \tag{A5}$$

注意，积分是针对 θ 所有可能值进行积分。然后，投资者求解如下问题：

$$\max_{\alpha} \int U\ (W_{T+H})\ p\ (R_{T+H} \mid z)\ \mathrm{d}R_{T+H} \tag{A6}$$

后验分布 $p\ (\theta \mid z)$ 与先验分布 $p\ (\theta)$ 成比例，θ 的似然函数为

$$p\ (\theta \mid z) \propto p\ (\theta)\ L\ (\theta;\ z) \tag{A7}$$

似然值反映了只用 z 的数据所给出的 θ 的最佳估计，先验分布是投资者在考察数据之前对 θ 的看法（例如，从理论上说，如果 θ 位于某个区域之外，那么这种看法就“不合理”）。

无参数不确定性：收益独立同分布

为了说明各种分布的计算，考虑股票收益独立同分布的最简单情形

$$r_t = \mu \mid \varepsilon_t \quad \varepsilon_t \sim iid\ (\mu,\ \sigma^2)\ .$$

我们如何计算（A4）的积分呢？以固定（已知）参数为条件的概率密度函数为 $p\ (R_{T+H} \mid z,\ \theta) \equiv p\ (R_{T+H} \mid r,\ \mu,\ \sigma)$，其中 $r_{T+j} = \mu + \varepsilon_{T+j}$ $(j = 1,\ 2,\ \cdots,\ H)$。因此，$R_{T+H} = r_{T+1} + {}_{T+2} + \cdots + {}_{T+H}$ 是均值为 $H\mu$ 方差为 $H\sigma^2$ 的正态分布。（A4）中的积分可以用数值模拟方法来进行计算。假设我们有某个函数 $g\ (y)$，希望计算

$$\int g\ (y)\ p\ (y)\ \mathrm{d}y$$

其中，$p\ (y)$ 是概率密度函数。积分可以用如下式子来进行逼近

$$(1/m) \sum_{i=1}^{m} g\ (y^{(i)})$$

其中，$y^{(1)}$，$\cdots$，$y^{(m)}$ 与概率密度函数 $p\ (y)$ 的选取独立。给定 μ 和 σ^2 的经验值，我们可以从 $N\ (\mu,\ \sigma^2)$ 中抽取 $m = 1$ 百万次（比方说）来得到 $R^{(i)}_{T+H} \sim (\mu H,\ H\sigma^2)$，然后对给定的 α 而言，我们可以用下式来计算（A4）中的和

$$EU\ (W_{T+H}) = \frac{1}{m} \sum_{i=1}^{m} \frac{\{\ (1-\alpha)\ \exp\ (r_f H)\ +\alpha \exp\ (r_f H + R^{(i)}_{T+H})\ \}^{1-\gamma}}{1-\gamma} \tag{A8}$$

接下来，我们在 $\alpha^{(j)} = \{0,\ 0.01,\ 0.02,\ \cdots,\ 0.98,\ 0.99\}$ 的格点值上进行搜索，从而确定 $\alpha^{(j)}$ 的最优值来使 $EU\ (W_{T+H})$ 达到最大。注意，我们对解 α 进限制，从而将卖空（$\alpha < 0$）、边际购买（$\alpha < 1$）等的可能性排除在外。（Barberis 和 Huang（2001）将 μ 和 σ^2 的后验均值用作这些变量的固定值，参见后面。）对不同的期限 H 而言，可以重复上述程序来考察最优的 α

值是如何随着期限的变化而变化的（尽管对独立同分布收益和无参数不确定性情形而言，我们期望 α 与期限独立，但这不必然是存在参数不确定性的情形）。类似地，针对风险规避参数 γ 的不同取值，我们可以计算出 α^*。

参数不确定性：收益独立同分布

我们必须计算（A6），这需要（A5）中的预测分布。为了计算后者，我们需要后验分布 $p(\theta \mid z)=p(\mu, \sigma^2 \mid r)$。在后验分布的构造过程中，我们能够使用的要么是信息先念（即 $\mu=\mu_0$的一些初始猜测），要么是扩散（无信息）先验，例如

$$p(\mu, \sigma^2) \propto 1/\sigma^2 \tag{A9}$$

对扩散先验而言，可以证明后验分布 $p(\mu, \sigma^2 \mid r)$ 为

$$\sigma^2 \mid r - IG\left(\frac{T-1}{2}, \frac{1}{2}\sum_{t=1}^{T}(r_t - \tilde{r}^2)\right) \tag{A10a}$$

$$\mu \mid \sigma^2, r \sim N(\tilde{r}, \sigma^2/T)$$

$$\tilde{r} = \sum_{t=1}^{T} r_t \tag{A10b}$$

当我们有扩散先念时，μ 的后验分布是极大似然估计值。方程（A10b）指出，μ 的最优估计值是样本平均值，估计的方差为 σ^2/T。如果投资者知道 σ^2，那么故事就到此为止。但是，如果投资者不知道 σ^2，那么就必须根据现有数据（$t=1, 2, \cdots, T$）来进行计算，（A10a）给出了计算方法，σ^2 的后验分布是一个逆 gamma（IG，Inverse Gamma）分布。如果样本是从任意时刻的后念分布中选择的（因此 T 和 $\tilde{r}$ 都是已知的），那么我们可以先从 IG 分布（A10a）中进行抽取来得到 $\sigma^{(i)}$（$i=1$），然后用这个数值从正态分布（A10b）中抽取 $\mu^{(i)}$。如果重复多次，$i=1, 2, \cdots, m$（为了获得足够的精度，m 要足够的大，比如说 1 百万次），那么我们得到了后验分布的精确表示。

假设 μ 的"真实"值为常数，那么由（A10b）可以看出，当 $n \to \infty$ 时，先验分布给出了 μ 的准确值（因为方差 σ^2/T 接近 0），尽管在我们的情形中 T 是固定的，它就是投资者进行买而持有决策时所能获得数据的长度。在信息先验情形中，后验分布变得更加复杂。然而，直觉告诉我们，θ 是后念值 θ_0 和样本估计值 $\hat{\theta}$ 的加权平均。

现在，我们就从后念概率密度函数 $p(\mu, \sigma^2 \mid r)$（$r=r_1, r_2, \cdots, r_T$）中对（μ, σ^2）进行了 $m=1$ 百万次抽取，每个数对（μ, σ^2）可以产生 $N(H\mu, H\sigma^2)$ 的一个样本点。这就给出了不确定情形中预测分布 $p(R_{T+H} \mid z)$（A5）的 1 百万个值，在 α 给定后，对这些值进行平均就得到（A8）中的 $EU(W_{T+H})$。对 $\alpha^{(j)}$ 进行格子搜索就可以得到（A8）中的最大值。

就月度数据而言，对由后念分布所得的（μ, σ^2）的 1 百万组值进行平均后就得到 $\hat{u}=0.0065$（标准差为 0.0039），σ^2 的平均估计为 0.0019（标准差为 0.0003）。均值的估计具有很大的误差，而 σ^2 的估计则要更准确一些（当我们从 $N(\mu, \sigma)$ 中进行抽取时，这些平均值可以用作 μ 和 σ^2 的"真实"值）。

重新调整

Barberis 和 Huang（2001）也重复了上述分析，但投资者每年都要对投资组合进行调整，因此他们的优化问题是

$$\max_{\alpha} E_{t_0}\left(\frac{W_k^{1-\gamma}}{1-\gamma}\right)$$

t_0 是决策期，每一次调整时都要向后移动 $k=12$ 个月。在每一个调整期中，投资者对参数不确定性的认知程度不同，他们的红利收益"初始值"也不同。求解这个问题是非常复杂的，需要使用贝尔曼方程和动态规划方法，可以用后向递归的方法来进行求解（Barberis 和 Huang（2001））。

第 17 章

理性泡沫与学习

目的

* 说明在求解欧拉方程时爆炸性理性泡沫和周期崩溃性泡沫是如何产生的，说明如何对“外生”泡沫的存在性进行检验。

* 说明泡沫如何依赖于基础变量（如红利），说明如何检验“内在泡沫”的存在性。

* 说明那些对红利或收益过程进行“递归学习”的模型如何导致价格更大的波动，即使理性定价公式（RVF）成立，模型的结论仍然是收益可预测。

* 说明“自我参考学习”（self - referential learning）模型如何收敛到理性预期均衡，说明为什么这些模型所生成价格的波动性比“纯”理性预期模型大。

17.1　理性泡沫

资产定价中“自促成泡沫”或“太阳黑子”的概念在市场建立之初就已经开始讨论。记录最早的（Garber（1990））著名泡沫包括 18 世纪 20 年代的南海股票泡沫和 17 世纪的郁金香泡沫。在“郁金香泡沫”中，从 1636 年 11 月到 1637 年 1 月，郁金香的价格急剧上升。1637 年 2 月，价格泡沫突然破裂，到 1639 年年底价格下降到最高值的 1%。20 世纪 20 年代股票价格上涨，突然在 1929 年崩溃。从 1994—2000 年，美国/英国的股票市场急剧上升，在 2000—2003 年突然崩溃。1982—1985 年期间美元即期汇率上升，在随后的几年中突然下降。这些都可以用“自促成泡沫”来进行解释。Keynes（1936）出名了，因为他发现股票价格不是受对基础变量的客观认识所主导，而是受“对平均观点进行预期的平均观点”所主导。他认为股票价格的预测类似于对选美竞赛中赢家的预测。客观的漂亮不一定是最重要的，重要的是一个裁判如何看待其他裁判对漂亮的看法，这在他们的投票模式中有所体现。

理性泡沫的产生是因为理性预期模型的解存在不确定性，主导股票价格的过程包含在欧拉方程中。今天你准备为某只股票所支付的价格取决于你所认为的未来股票价格，这个价格在未来的某个时刻能给你带来收益。但你所认为的价格又取决于未来更远时刻的预期价格。欧拉方程决定了价格序列，但不能给出唯一的价格水平，除非随意给定，我们可以施加终端条件（如横截性条件）来得到唯一解。然而，一般而言，欧拉方程不能排除价格包含爆发性泡沫的可能性。（这一点有待斟酌，尤其是在 Tirole（1985）的代表性投资者模型中，当寿命无穷大的理性交易者数目为有限数时模型的解是唯一的，Tirole（1985b）表明，只有当经济增长率比稳定状态下的资本收益高时，才有可能产生泡沫。）

人们当然可以对股票价格的持续上升或下降试图进行解释，认为可能是因为某些非理性

行为（如羊群效应、市场心理）的存在，但最新研究强调这种急剧运动或泡沫可能与理性行为假设一致。尽管交易者完全理性，真实的股票价格也可能包含泡沫因素，因此股票价格与基础价值之间存在差异。

欧拉方程和理性定价公式

我们考察在投资者同质且理性、市场信息有效的情形中股票的市场价格如何偏离基础价值。为此，我们证明股票的市场价格等于股票的基础价值加上一个“泡沫项”，在没有超额收益的情况下理性投资者仍然愿意持有股票。为了简化推导，假设：（i）代理人属于风险中性，具有理性预期；（ii）投资者的必要（实际）资产收益为常数 $E_tR_{t+i}=k$。欧拉方程为：

$$P_t=\delta\ (E_tR_{t+1}+E_tD_{t+1}) \tag{1}$$

其中，$\delta=1/(1+k)$。在理性预期下，向前反复替代就可以得到解

$$P_t=P_t^f=\sum_{i=1}^{\infty}\delta^iE_tD_{t+i} \tag{2}$$

我们在这里假设横截性条件成立（即 $\lim\limits_{n\to\infty}(\delta^nE_tD_{t+N})=0$），横截性条件能保证由（2）给出的价格是唯一的，我们用基础价值 P_t^f 来表示这一价格。理性泡沫背后的基本思想是，对 P_t 而言存在满足欧拉方程的另外一个数学表达式，即

$$P_t=\sum_{i=1}^{\infty}\delta^iE_tD_{t+i}+B_t=P_t^f+B_t \tag{3}$$

B_t 就是“理性泡沫”。因此，B_t 就是实际市场价格 P_t 与基础价值 P_t^f 偏离程度的度量。到目前为止，我们还没有说明 B_t 的性质。显然，如果 B_t 相对于基础价值而言很大，那么实际价格就会极大地偏离它的基础价值。

要使（3）满足（1），我们必须对 B_t 的动态行为施加一些限制性条件，我们通过构造潜在的矛盾来确定这些限制条件。首先假设（3）是（1）的一个有效解，这就为 B_t 的动态行为施加了限制。在时刻计算（3）在 $t+1$ 时刻的数学期望

$$\begin{aligned}E_tP_{t+1}&=E_t\ [\delta E_{t+1}D_{t+2}+\delta^2E_{t+1}D_{t+3}+\cdots+B_{t+1}]\\&=[\delta E_tD_{t+2}+\delta^2E_tD_{t+3}+\cdots+E_tB_{t+1}]\end{aligned} \tag{4}$$

这里用到了迭代数学期望公式 $E_t\ (E_{t+1}D_{t+j})=E_tD_{t+j}$。欧拉方程（1）的右边包含 $\delta\ (E_tP_{t+1}+E_tD_{t+1})$ 这一项，使用（4）就可以看到

$$\delta\ [E_tD_{t+1}+E_tP_{t+1}]=\delta E_tD_{t+1}+[\delta^2E_tD_{t+2}+\delta^3E_tD_{t+3}+\cdots+\delta E_tB_{t+1}] \tag{5}$$

将（2）中 P_t^f 的定义代入（5）的右边就得到

$$\delta\ [E_tD_{t+1}+E_tP_{t+1}]=P_t^f+\delta E_tB_{t+1} \tag{6}$$

将（6）代入（1），

$$P_t=P_t^f+\delta E_tB_{t+1} \tag{7}$$

但是，我们现在就得到了一个矛盾，因为（3）和（7）一般而言不可能同时是（1）的解。如果

$$E_tB_{t+1}=B_t/\delta=(1+k)\ B_t \tag{8}$$

那么我们可以将（3）和（7）两个解视为等价。于是，（3）和（7）不可能是满足（1）的同一表达式。多种方法都证明，泡沫必须是鞅，由（1）就可以注意到这一点，我们在 $B_t=E_tB_{t+1}$ 的假设下就可以得到 $P_t+B_t=\delta\ (E_tP_{t+1}+E_tD_{t+1}+E_tB_{t+1})$。更为一般的情形是，（8）意味着

$$E_tB_{t+m}=B_t/\delta^m \tag{9}$$

因此（除了已知的贴现因子外）B_t 必须是一个鞅：对未来泡沫值的最优预测依赖于它的当前值。当泡沫解满足欧拉方程时，它违背了横截性条件（因为 $B_t\neq0$），因为 B_t 是任意的，所以（3）中的股票价格不是唯一的。

这个数学等式表示的是哪种类型的泡沫呢？注意，由（8）有 $E\ (B_{t+1}/B_t)-1=k$，即泡

沫的预期增长率等于投资者愿意持有股票的必要收益率，这时的泡沫是一个有效解。因为实际市场价格中的泡沫部分按照必要收益率进行支付，所以投资者并不在乎是否为泡沫进行支付（而不是为基础价值进行支付）。泡沫是一种自促成预期。

考虑一个简单的例子，预期红利和时刻的泡沫价值 B_t 均为常数，$B_t=b$（>0）。此时，泡沫是可以确定的，并以 k 的速度进行增长，因此，$E_tB_{t+m}=(1+k)^m b$。于是，只要泡沫存在，即使红利为常数，根据（3），$t+m$ 时刻的股票价格为

$$P_{t+m}=\frac{\delta D}{(1-\delta)}+b\ (1+k)^m \tag{10}$$

即使基础变量（如红利）能说明实际股票价格应为常数，泡沫的存在还是意味着实际价格可能持续增长，因为 $(1+k)>1$。

在上述例子中，因为泡沫在进行增长而基础价值却保持不变，所以股票价格中泡沫的比例会不断增加。事实上，即使红利不是常数，股票价格也总是以低于泡沫增长率（k）的速度进行增长，因为红利支付为

$$(E_tP_{t+1}/P_t)-1=k-E_tD_{t+1}/P_t \tag{11}$$

在泡沫存在的情形中，投资者仍然使用一切可获得的信息来预测股票价格和收益率。因此，预测误差与 t 时刻的信息独立，超额收益无法预测。因此，在检验泡沫时，信息有效性的检验没有什么作用。然而，泡沫不允许（超额）利润存在，因为与未来红利和泡沫有关的所有信息都已经被当前价格所吸收：泡沫满足公平游戏规则。

我们的泡沫模型可以扩展到包括泡沫以 $(1-\pi)$ 的概率破裂或以 π 的概率持续下去的情形（Blanchard（1979））。

$$B_{t+1}=B_t\ (\delta\pi)^{-1}\quad（概率为\ \pi） \tag{12a}$$

$$=0\quad（概率为\ 1-\pi） \tag{12b}$$

这一结构也满足鞅性。值得注意的是，这些理性泡沫模型没有告诉我们泡沫是怎样开始以及怎样结束的，只是告诉我们泡沫一旦开始后的时间序列特性。泡沫对基础变量模型和价格通常的理性定价公式而言是外生的。

我们在上面注意到，投资者无法区分价格的上涨是由基础变量导致的还是因为泡沫存在。只要泡沫部分在下一期能给他们带来持久的必要收益，个体就不介意支付超过基础价值的价格。理性泡沫的一个隐含条件是它们不可能为负数（即 $B_t<0$）。这是因为泡沫部分比股票价格下降得快，因此负的理性泡沫最终会以 0 价格结束。理性代理人能够意识到这一点，知道它们最终会破裂。但是，运用反向归纳法，泡沫必须即刻破裂，因为没有人愿意为早期的“泡沫溢价”进行支付。因此，如果实际价格 P_t 比基础价值 P_t^f 低，那么这不可能是因为理性泡沫的存在而引起的。如果泡沫不可能为负数，那么一旦泡沫价格为 0，它就不可能形成新的泡沫。这是因为在理性泡沫中创新（innovation）$(B_{t+1}-E_tB_{t+1})$ 的均值为 0。如果泡沫重新形成的话，那么创新的均值不可能为 0，因为泡沫可能朝一个方向变动（比如说增加）以便重新形成泡沫。

理论上讲，因为股票价格没有上限，所以正的泡沫是可能的。然而，在这种情形中，可能出现这种情况：由于泡沫部分在实际价格中的比例增加，价格中的基础价值比例变得相当的小。人们可以推测，这可能意味着个体会在未来的某个时刻感觉到泡沫会破裂。一旦投资者认为泡沫在未来某个时刻破裂的话（无论是什么原因），它现在就会破裂。为了看清这一点，假设个体认为泡沫在 2030 年破裂，他们必定意识到 2029 年的市场价格仅仅是基础价值的反应，因为人们预期泡沫在第二年会破裂。但是，如果 2029 年的股票价格仅仅反映基础价值的话，那么由反向归纳法可知，在所有更早时期中的价格也是这个样子。因此，现在的价格也只反映了基础价值。似乎在现实世界里只有当个体的投资期限比预期泡沫破裂的时间短时，理性泡沫才会存在。也就是说，人们愿意支付比基础价值高的价格，因为他们相信有人愿意在将来支付更高的价格。因此，投资者没有远见（近视），在未来某个时刻 $t+N$，股票价格依赖于他们怎

么看待价格和其他投资者如何看待价格。

17.2 理性泡沫的检验

很容易证明，Shiller 方差边界不等式的违背并不意味着理性泡沫存在。从直觉来看，这是因为在检验中用到的最终价格 P_N 包含某些泡沫成分。更为正式的表述是，在计算完全预见价格时，要使用 RVF 中无限期贴现值的逼近技巧，所计算的价格为 P_t^*

$$P_t^* = \sum_{i=1}^{N} \delta^i D_{t+i} + \delta^N P_{t+N} \tag{13}$$

这里的 P_{t+N} 是数据集合在最后时刻的实际市场价格。在必要（实际）收益为常数的零假设下，方差边界不等式为 Var（P_t）≤Var（P_t^*）。然而，泡沫已经嵌入到零假设之中。为了看清这一点，注意，当理性泡沫存在时

$$P_t = P_t^f + B_t \tag{14}$$

且 $E_t(B_{t+N}) = (1+k)^N B_t = \delta^{-N} B_t$。如果我们现在用一个包含泡沫的项 $P_{t+N} = P_{t+N}^f + B_{t+N}$ 来代替（13）中的 P_N，那么

$$E_t P_t^* = P_t^f + \delta^N E_t B_{t+N} = P_t^f + B_t \tag{15}$$

因此，即使在泡沫存在的情形中，由（14）和（15）我们也有 $P_t = E_t P_t^*$。对泡沫的早期检验（Flood 和 Garber（1980））假设存在随机泡沫，即 $P_t = P_t^f + (B_0/\delta^1)$，这里的 B_0 是样本期初始时刻的泡沫价值。因此，在进行回归时，就多了一项 $(B_0/\delta)^t$。当 δ 已知时，泡沫存在性检验的零假设为 $B_0 \neq 0$。不幸的是，因为 $1/\delta > 1$，所以回归系数 $(1/\delta)^t$ 发散。这意味着对 B_0 的检验依赖于非标准分布（non - standard distributions），正确推断也可能是一个问题（具体细节请参见 Flood、Garber 和 Scott（1984））。

West（1987a）提供了一个更为巧妙的泡沫检验方法，检验涉及用两种方法来计算一个特殊的参数。在不存在泡沫的假设下，这两个参数的估计值在一定的统计精度范围内必须相等；当理性泡沫存在时，它们就不相等。这种方法的好处是它不需对泡沫过程进行特殊的参数化处理：与红利有关的任何泡沫都可以检测出来。为了说明这种方法，首先用（工具变量）“可观测的”欧拉方程的估计值来对 δ 进行估计

$$P_t = \delta(P_{t+1} + D_{t+1}) + \mu_{t+1} \tag{16}$$

其中，$\mu_{t+1} = -\delta[(P_{t+1} + D_{t+1}) - E_t(P_t + D_{t+1})]$。现在假设红利服从一阶自回归过程（$AR$（1））

$$D_t = \alpha D_{t-1} + v_t \quad (|\alpha| < 1) \tag{17}$$

在无泡沫的假设下，由 RVF 和（17）就得到

$$P_t = \Psi D_t + \varepsilon_t \tag{18}$$

其中，$\Psi = \delta\alpha/(1 - \delta\alpha)$，$\varepsilon_t$ 的产生是因为计量经济学家所拥有的信息集合只是真实信息集合的一个子集。用 $\hat{\Psi}$ 表示 Ψ 的间接估计值，它可以通过（16）的 δ 回归估计值和（17）的 α 来得到。然而，Ψ 的直接估计值 $\hat{\Psi}^*$ 可以通过（18）P_t 对 D_t 进行回归来得到。在无泡沫的零假设下，Ψ 的直接估计值和间接估计值应相等。

考虑这样的情形，当泡沫存在且 $P_t = P_t^f + B_t = \Psi D_t + B_t$ 时，P_t 对 D_t 的回归方程中包含一个被遗漏的变量，即泡沫，Ψ 的估计值 $\hat{\Psi}^r$ 将是不一致估计

$$\text{plim}\,\hat{\Psi}^r = \Psi + \text{plim}\,(T^{-1} \sum D_t^2)^{-1}\,\text{plim}\,(T^{-1} \sum D_t B_t) \tag{19}$$

如果泡沫 B_t 与红利相关，那么 $\hat{\Psi}^r$ 是有偏的（如果 Cov（D_t，B_t）>0，那么偏斜是向上的），且是不一致的。但欧拉方程和红利预测方程仍然给出参数和 $\hat{\Psi}$ 的一致估计值。因此，当泡沫

存在时，$\hat{\Psi} \neq \hat{\Psi}^r$（Hausman（1978）的检验可以用来对系数的可能变化进行检验）。

West（1987a）使用了上述方法，数据包括 Shiller（1981）中1871—1980年的指数数据（以及1928—1978年的道·琼斯指数）。West 发现两组估计值之间存在较大的差异，因此要拒绝无泡沫的零假设。然而，这个结果可能的原因是，均衡收益模型或红利行为模型是错误的。事实上，West 认识到了这一点。他还发现，当红利服从 ARMA 过程时结果具有合理的稳健性。与此相对照的是，当贴现率具有时变性时，没有理由拒绝不存在泡沫的零假设。Flood、Hodrick 和 Kaplan（1994）指出，如果对欧拉方程进行两期迭代，待估的（“两期”）欧拉方程没有很好地进行设定，那么 Λ_t 的估计值将是有偏的。因为 RVF 的推导需要对欧拉方程进行无限次迭代，这就使得人们对 West（1987a）中的估计值产生了怀疑。

West（1988a）对泡沫检验方法进行了深入发展，这种方法再一次将基于不同信息集合的两个估计结果进行比较。第一个集合是受限信息集合 Λ_t，它包含当前和历史红利信息；另一个信息集合是对未来红利的最优预测，即市场价格 P_t。在不存在泡沫的零假设下，基于受限信息集合 Λ_t 的预测会产生更大的误差（严格地说，创新的方差），但 West 得到的结果却相反。这个证据对无泡沫假设进行了反驳，当然，这不必然与投资者的偏好不一致。

对理性泡沫存在性的一些检验是建立在对红利和价格数据的平稳性进行考察的基础上的。外生泡沫给价格带来了一个发散的部分，它不会出现在基础变量（即红利或贴现率）之中。因此，如果股票价格 P_t 的增长速度比 D_t 快，那么这有可能是由泡沫项 B_t 的存在所引起的。这些直观的概念可以通过单位根或协整来进行表述。当使用 RVF 时（在贴现率为常数的假设下），如果 P_t 和 D_t 是单位根过程，那么它们应该具有协整关系。

如果红利（或红利的对数值）是一阶自积序列，I（1），$P_t = [\delta/(1-\delta)]D_t$，那么 P_t 一定是 I（1）序列，且与 D_t 具有协整关系，$z_t = P_t - \delta/(1-\delta)D_t$ 是 I（0）平稳序列。Diba 和 Grossman（1988）使用总的股票价格和红利指数后发现，ΔP_t 和 ΔD_t 是平稳变量，P_t 和 D_t 具有协整关系，因此应拒绝（8）所表示的爆发性泡沫的存在性。

不幸的是，当存在 Evans（1991）所谓的周期性泡沫时，上述检验的解释可能具有误导性。Evans 所考察的理性泡沫类型是始终为正的泡沫，泡沫可能会“破裂”，泡沫在破裂（回到正的均值）之前增长速度很快。周期性泡沫的轨迹与持续增长性泡沫的轨迹不同（见图1）。

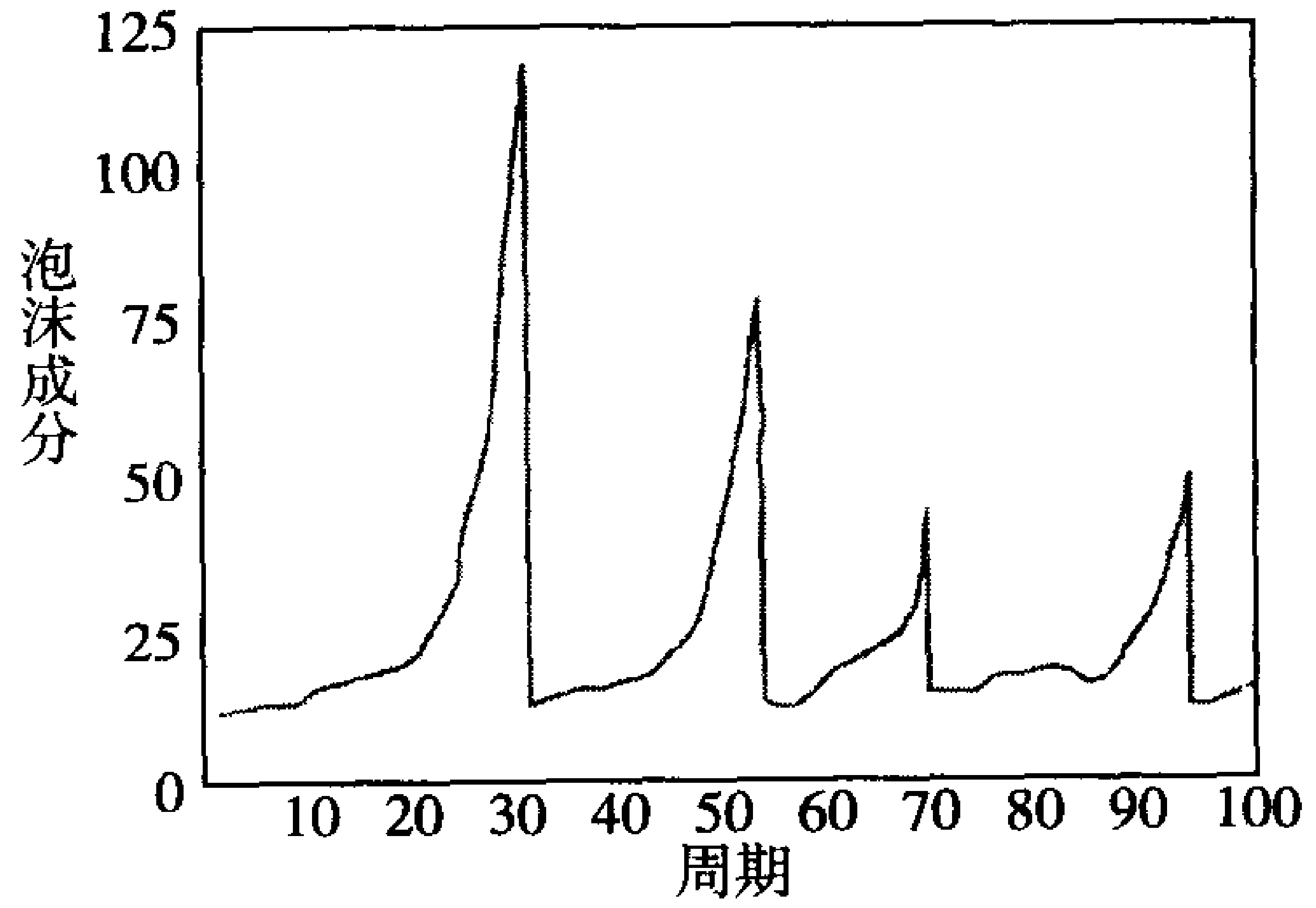


图1　泡沫成分。资料来源：Evans（1991）。**美国经济协会授权。**

人们可以从直觉上看出为什么对 P_t 是否为非平稳序列（I（1））进行考察不能检测出图1所示的泡沫。（Dickey - Fuller）平稳性检验的本质是试图考察时间序列是否具有很强的趋势性，或它们的无条件方差是否是一个常数。显然，图1中的泡沫没有很强的上升趋势，尽管方差随着时间是变化的，如果（在任何给定的时期内）泡沫破裂的概率很高，那么对泡沫的检验尤其困难。如果泡沫破裂的概率很低，那么我们就接近于 Diba 和 Grossman 所考察过的“爆发性泡沫”（$E_t B_{t+1} = B_t/\delta$），我们在这里预期平稳性标准检验会得到更加丰富的结果。

为了更具启发性（也为了进一步简化问题），Evans 使用 MCS 来生成周期性泡沫的一个时间序列。将泡沫项和基础价值 P_t^f 相加（例如，在 D_t 是带漂移项的随机游走的假设下）就得到股票价格的生成序列，可以对这个价格序列进行单位根存在性的标准检验。他发现单位根检验的结果依赖于泡沫不崩溃的概率 π。当 $\pi<0.75$ 时，超过 90% 的模拟值错误地表明 ΔP_t 是平稳序列且 P_t 和 D_t 具有协整关系。因此，“周期崩溃性泡沫”不能用标准单位根检验来进行检测。（原因是，标准单位根检验假设的是线性自回归过程，而 Evans 模拟涉及的是复杂的非线性泡沫过程）。因此，Diba 和 Grossman 未能检验出股票价格中持续爆发性泡沫的存在并没有排除其他类型的理性泡沫不存在。显然，为了对周期崩溃性泡沫进行检验，我们需要更加复杂的非平稳性统计检验方法（如 Hamilton（1994））。

17.3 内在泡沫

到目前为止所讨论的泡沫类型存在的一个问题是：泡沫是外生的（相对于基础变量（如红利）来说是外生的）。泡沫项是股票价格欧拉方程的另外一个解（严格地说，是解的共同部分（homogeneous part））。Froot 和 Obstfeld（1991）提出了另外一种类型的泡沫，他们称之为内在泡沫。用“内在”这个术语是因为泡沫（以非线性的方式）依赖于基础变量，如（实际）红利水平。因此如果基础变量是常数，那么泡沫项也是常数。但是，也可能出现的情形是，泡沫项会随着红利水平的上升（下降）而上升（下降）。就内在泡沫的这种形式而言，如果红利具有持续性，那么泡沫也具有持续性，股票价格与基础价值之间存在持久的偏离。另外，内在泡沫能使股票价格对红利（基础变量）变化产生过度反应，这与经验证据一致。

为了分析这种形式的内在泡沫，假设必要收益率为常数 r（在连续时间情形中）。欧拉方程为

$$P_t=\mathrm{e}^{-r}E_t\ (D_{t+1}+P_{t+1}) \tag{20}$$

这意味着基础价值 P_t^f 为（假设横截性条件成立）

$$P_t^f=\sum_{k=1}^{\infty}\mathrm{e}^{-r(k)}E_t\ (D_{t+k}) \tag{21}$$

然而，如果 B_t 是鞅，$B_t=\mathrm{e}^{-r}\ [E_tB_{t+1}]$，那么 $P_t=P_t^f+B_t$ 仍然是欧拉方程的一个解。寻找泡沫与红利之间的一个非线性函数并使 B_t 为鞅且满足欧拉方程，我们就能构造出“内在泡沫”：

$$B\ (D_t)\ =cD_t^{\lambda}\qquad (c>0,\ \lambda>1) \tag{22}$$

如果红利的对数值服从随机游走，漂移项为 μ，条件方差为 σ^2，即 $\ln\ (D_{t+1})\ =\mu+\ln\ (D_t)+\varepsilon_{t+1}$，那么泡沫解 $\hat{P}_t$ 为

$$\hat{P}_t=P_t^f+B\ (D_t)\ =\alpha D_t+cD_t^{\lambda} \tag{23}$$

其中，$\alpha=\ (\mathrm{e}^r-\mathrm{e}^{\mu+\sigma^2/2})^{-1}$。基本价值解 $P_t^f=\alpha D_t$ 是 Gordon（1962）增长模型的随机版本，Gordon 模型给出的确定性条件下的解为 $P_t^f=\ (\mathrm{e}^r-\mathrm{e}^{\mu})^{-1}$。显然，由（23）可以看出，与“基础价值”解相比（$\partial P_t^f/\partial D_t=\alpha$），由于泡沫的存在（$\partial P_t/\partial D_t=\alpha+c\lambda D_t^{\lambda-1}$），股票价格对当前红利存在过度反应。Froot 和 Obstfeld 通过（$r,\ \mu,\ \sigma^2$）的合理值来模拟（22）中的内在泡沫，估计 c 和 λ 的值（见下面），得到 ε_{t+1} 的独立同分布。他们比较了纯基础价值的轨迹、由（23）给出的随机泡沫 $\hat{P}$ 的轨迹和同时依赖红利和时间的内在泡沫 $\hat{P}_t$ 的轨迹。$\tilde{P}_t$ 由下式给出

$$\tilde{P}_t=aD_t+bD_t\mathrm{e}^{(r-\mu-\sigma^2/2)t} \tag{24}$$

（24）中的内在泡沫依赖于时间，允许我们对带参数的泡沫检验进行比较，后者通常涉及一个确定的指数时间趋势（Flood 和 Garber（1990），Blanchard 和 Watson（1982），Flood 和 Garber（1994））。图 2 给出了这三种价格序列的模拟值。显然，内在泡沫使股票价格 $\hat{P}_t$ 的轨

迹看起来很合理，$\hat{P}_t$ 一直位于基础价值 P_t^f 之上（尽管在其他模拟中，内在泡沫 $\hat{P}_t$ 可以位于基础价值 P_t^f 的上方，但接下来会朝 P_t^f 方向“破灭”）。由图 2 我们可以看出，在有限样本中，内在泡沫看上去不是爆发性泡沫，因此很难用统计检验进行检测。（另一方面，时间依赖型内在泡沫 $\tilde{P}_t$ 的轨迹看上去具有发散性，标准检验更能揭示这一点）。

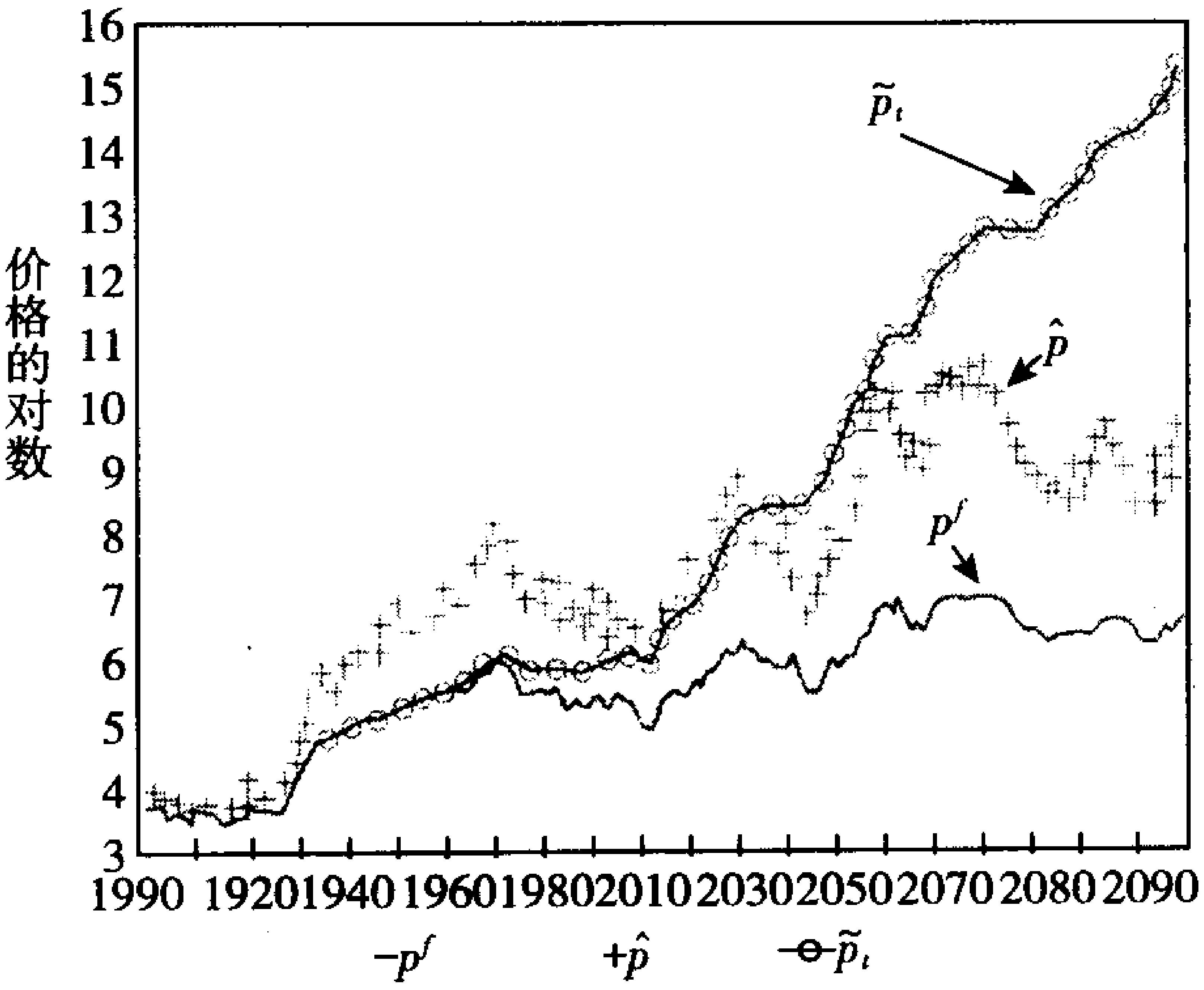


图 2　模拟的股票价格轨迹

资料来源：Froot 和 Obstfeld（1991）。美国经济协会授权。

Froot 和 Obstfeld 在对（23）进行简单变换后来对内在泡沫的存在性进行检验。

$$P_t/D_t = c_0 + cD_t^{\lambda-1} + \eta^t \tag{25}$$

其中，不存在内在泡沫的零假设为 H_0: $c_0 = \alpha$ 且 $c = 0$（$\alpha = e^r - e^{\mu+\sigma^2/2}$）。对实际标准普尔指数而言，具有代表性的值为 $e^r = 1.09$（p. a）（p. a 就是每年的意思），对实际红利过程而言，$\mu = 0.011$，$\sigma = 0.122$，因此的样本均值等于 14。在无泡沫的零假设下，P_t 和 D_t 具有协整关系，协整参数为 $c_0 = \alpha = 14$。在 P_t 对 D_t 的简单 OLS 协整回归中，Froot 和 Obstfeld 发现 $P_t = \Psi + 37D_t$，所以 P_t 对 D_t 存在过度反应。另外，$P_t - 14D_t$ 是非平稳的，P_t 和 D_t 不存在协整关系。“仅含基础价值的”解 $P_t = \alpha D_t$ 仍然意味着 $\ln P_t$ 和 $\ln D_t$ 是协整的，且协整参数为 1。然而，估计结果表明参数值落在 1.6 ~ 1.8 的范围之内，（$\ln P_t - \ln D_t$）可能不是平稳的。因此，从表面值来看，这种检验倾向于拒绝（无泡沫的）基础变量模型。然而，Froot 和 Obstfeld 注意到，OLS 协整参数可能存在严重的偏差（Banerjee、Dolado、Galbraith 和 Hendry（1993）），这些检验的功效值得怀疑。

Froot 和 Obstfeld 在对（25）进行估计的基础上考虑内在泡沫存在性的直接检验。代表性结果是

$$(P/D)_t = \underset{(2.28)}{14.6} + \underset{(0.12)}{0.04}D_t^{1.6(1.1)} \tag{26}$$

1900—1988 年期间的度数据，$R^2 = 0.57$，（·）表示 Newy - West 标准差。

尽管在对（25）中 $c = 0$ 和 $\lambda - 1 = 0$ 进行检验时还存在一些小样本计量经济问题，上述结果还是能部分地支持内在泡沫的存在性。$c = 0$ 和 $\lambda - 1 = 0$ 的联合零假设遭到了拒绝。然而，因为我们没有拒绝 $c = 0$ 的零假设，所以实证结果不具有决定性。人们可以模拟基础价值 $P_t^f = 14.6D_t$ 和（23）所给出的包含内在泡沫的价格 $\hat{P}_t$，并将两个序列进行比较。内在泡沫的轨迹（图 3）$\hat{P}_t$ 比 P_t^f 更接近于真实的股票价格 P_t 的轨迹。第二次世界大战后，泡沫的规模很大。

事实上，最后时刻 S&P 指数中泡沫成分非常巨大。

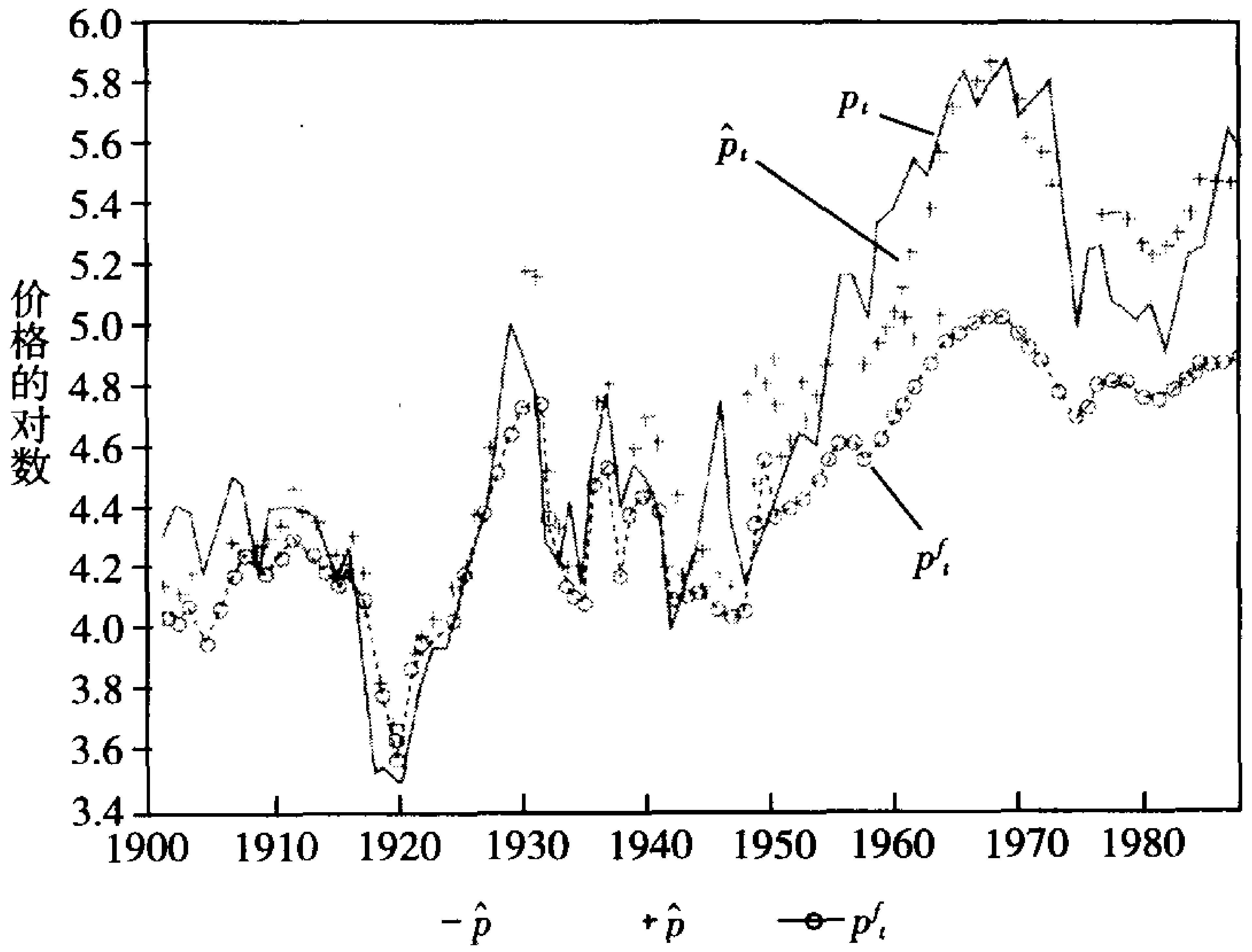


图 3　实际股票价格和预测股票价格

资料来源：Froot 和 Obstfeld（1991）。美国经济协会授权。

最后，Froot 和 Obstfeld 考察了结果对不同红利模型（使用蒙特卡罗模拟方法）和 D_t 的不同函数形式或回归方程（22）中其他定性时间趋势的敏感性。基础内在泡沫方程（23）的估计非常稳健。

Driffell 和 Sola（1998）对 Froot - Obstfeld 模型进行了重复并假设红利增长经历了区制（状态）转移，特别地，他们假设红利的条件方差随着样本的不同而发生变化。美国（实际）红利增长的图形表明，1900—1920 年期间的方差相对较小，1920—1950 期间的方差变化较大，1950 年以后的方差相对较小且接近于常数。Driffell 和 Sola 用 Hamilton（1989）的两状态马尔科夫区制转移模型来对红利增长过程进行建模，这证实了图形中的结果。相应于两个状态（“高”方差和“低”方差），他们得到了形如（23）的两个方程。然而，具有内在泡沫的价格图形与 Froot 和 Obstfeld 中的图形（图（3）中的 $\hat{P}_t$）非常相似，因此，这种特殊形式的变体似乎没有造成太大的差异。

Froot 和 Obstfeld 方法的统计推断需要很多统计假设（一些假设是我们已经提到过的假设）。因为这个原因，他们会很满意地说：“上述结果只是表明，存在与泡沫一致的情形”。他们可能会同意如下结论：证据也与其他假设一致。

计量经济问题

在理性泡沫的检验过程中存在一些严重的计量经济问题，对结果的解释也存在问题。所产生的计量经济问题包括使用有限数据集合上可能的非平稳序列所进行的分析、存在发散回归系数时渐进检验统计量的行为、获得非线性参数精确估计值的标准问题（内在泡沫）、对异方差和移动平均误差的修正问题。对理性泡沫的检验经常取决于“均衡的期望收益模型是否是正确的模型”，因此，我们是对联合假设进行检验。拒绝无泡沫的零假设可能只是说明基于基础变量的模型不是正确的模型。

解释理性泡沫检验结果的另外一个困难是所谓的 Peso 问题，也就是某种形式的“省略变量”问题。假设市场上的投资者在所考察的样本期内拥有信息，未来红利可能快速增长，但在实际数据中，事件不会发生，红利以“正常速度”增长。在样本数据中，股票价格可能大幅

上升，但红利只有适度的上升。股票价格看上去似乎对（当前）红利存在过度反应，更为重要的是，这种价格上升可能被错误地解释为泡沫。当价格变动非常大的时候（如上世纪 90 年代的长期牛市，及随后的 2000—2003 崩溃），解释方面的问题会变得非常尖锐。当我们研究周期崩溃性泡沫时，使用标准检验方法来检测这种现象似乎不可能。

17.4　学习

在这一节里，我们不介绍有关动态学习和理性预期均衡收敛的大量文献（参见 Evans 和 Honkapohja（2001）），而是把注意力集中在强调过度波动性和可预测性的学习方法上。我们的“基准模型”仍然是纯理性预期模型，在纯理性预期模型中，代理人知道“真实的”经济结构（至多再有一个白噪声误差过程）。然而，我们将标准理性预期模型和简单学习过程所得到的结果进行比较。在这些学习模型中，代理人是“理性的”但不是理性预期的，他们不知道红利和收益方程中参数的真实值，他们必须使用有限数据来估计这些参数。我们假设，一旦获得更多的数据，代理人就对总体参数的估计值进行最优化更新。然后，他们使用 RVF 中的弱理性预测（基于受限信息集合）来获得“学习机制下的价格”估计值，然后将这个估计值与所观察到的价格（和股票收益）进行比较来考察“学习模型”对真实数据的模拟是否比“纯”理性预期模型更好。我们在后面的章节中讨论一个补充问题：学习对投资组合选择的影响。

Gordon 增长模型

一个简单而又直观的学习模型是 Barsky 和 DeLong（1993），他们使用简单的 Gordon（1962）增长模型，其中（实际）股票价格为 $P_t = D_{t+1}/(P-g)$，R 是（已知的）必要收益，g_t 是基于 t 时刻信息所预期的“稳定状态”红利增长率。Barsky 和 DeLong（1993）用这个模型来计算基础价值 $V_t = D_{t+1}/(R-g_t)$，假设代理人不断更新未来红利增长率的估计值。然后，他们将 V_t 和 1880—1980 年期间实际 S&P500 指数 P_t 进行比较。

对 $(R-g)^{-1}$ 为常数 20 的最简单情形而言，平均 P/D 比率也等于 20，$V_t = 20D_{t+1}$。10 年期 V_t 的变动幅度高达 P_t 方差的 67%（参见 Barsky - DeLong（1993），Table III，p. 302），在第二次世界大战前这两个数据序列的运动非常接近。

他们为红利增长所建议的简单外生学习过程是熟悉的“适应性预期”：

$$E_t\Delta d_{t+1} = \theta E_{t-1}\Delta d_t + (1-\theta)\Delta d_t \quad (0<\theta<1) \tag{27}$$

或

$$E_t\Delta d_{t+1} - E_{t-1}\Delta d_t = (1-\theta)[\Delta d_t - E_{t-1}\Delta d_t] \tag{28}$$

$t-1$ 到 t 之间预期的变化依赖于前期的预测误差，调整参数为 $(1-\theta)$。如果 θ 很大（如 0.95），那么红利增长的新信息对预期变化的影响不大。可以证明，如果红利增长 Δd_{t+1} 遵循如下 IMA（1，1）过程，那么具有适应性预期的预测计划就是理性的：

$$\Delta d_{t+1} = \Delta d_t + (\varepsilon_{t+1} - \theta\varepsilon_t) \tag{29}$$

为了看清这一点，对（29）求数学期望并滞后一期，

$$E_t\Delta d_{t+1} = \Delta d_t - \theta\varepsilon_t \tag{30a}$$

$$E_{t-1}\Delta d_t = \Delta d_{t-1} - \theta\varepsilon_{t-1} \tag{30b}$$

由（30a）和（30b）可以得到

$$E_t\Delta d_{t+1} - \theta E_{t-1}\Delta d_t = \Delta d_t - \theta\Delta d_{t-1} - \theta(1-\theta L)\varepsilon_t \tag{31}$$

由（29）我们有 $\Delta d_t - \Delta d_{t-1} = (1-\theta L)\varepsilon_t$，在（31）中使用这个方程就得到（28）

$$E_t\Delta d_{t-1} - \theta E_{t-1}\Delta d_t = (1-\theta)\Delta d_t \tag{32}$$

往回代入，(27) 可以写成

$$E_t\Delta d_{t+1} = (1-\theta)\sum_{i=0}^{t-1}\theta^i\Delta d_{t-1} + \theta^i\Delta d_0 \tag{33}$$

（30a）减去（30b）并使用 IMA 过程（29），我们就得到另外一种形式的适应性预期

$$E_t\Delta d_{t+1} = E_{t-1}\Delta d_t + (1-\theta)\ \varepsilon_t \tag{34}$$

方程（27）、(32)、(33）和（34）都是等价表示。方程（33）表明，Δd_{t+1}的一步预测依赖于历史红利增长的长期（几何）分布（因此，预测值的变化非常缓慢）。方程（34）讲述的是一个类似的故事，因为当 $\theta \approx 0.95$ 时，对红利增长预测值的修正很少依赖新的信息 ε_t（类似于方程（28））。Barsky - DeLong（1993）注意到，对（29）中 IMA（1，1）过程进行估计所给出的 θ 值落在 0.95 ~ 1 的范围之内，即使在大样本中，我们也不能说 $1-\theta$ 是一个很小的正数或者就是 0。

他们使用（28）来更新每一期的红利增长预测值 $E_t\Delta d_{t+1}$，假设这个值能模拟代理人眼中的“稳定状态”增长率 g_t。改变 g_t 就可以得到关于 V_t 的一个序列（$V_t = D_{t+1}/(R-g_t)$），它比假定 g 为常数时的序列有着更强的波动性。他们发现 10 年中 V_t 的变化达到 P_t 波动率的 76%。(参见 Barsky 和 DeLong（1993）表 III，*p.* 302)。然而，也应注意的是，尽管这个模型能部分地解释 P_t 的长期波动，但它仍不能很好地解释短期（比如一年，甚至 5 年内）波动，也不能很好地解释（平稳的）价格/红利比的运动。因此，上述证据至多和“长期中实际红利与实际价格一起运动（因此，可能存在协整关系）”的观点一致，但是，即使模型包含红利增长的学习过程，股票价格和基础价值在多年的时期中也存在相当大的差异。

常数贴现率

Timmermann（1993）使用 MCS 考察了学习对价格波动和可预测性的影响，并将这些结果与 *RE* 模型进行了比较（两个模型中的贴现率都为常数）。学习模型与 Barske - DeLong 模型很相似，但 MCS 允许我们判断结果的可能分布而不是只给出一个点估计值。假设预期红利增长率为常数

$$\ln D_t = \mu + \ln D_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim niid\ (0,\ \sigma^2) \tag{35}$$

因此，在具有理性预期的标准常数收益率（贴现率）模型中，股票价格为

$$P_t^{\mathrm{RE}} = \frac{1+g}{R-g}D_t = kD_t \tag{36}$$

其中，$(1+g) = \exp\ (\mu + \sigma^2/2)$。在 RE 下，价格的波动率与红利波动率成比例，使用红利/价格比无法对预期收益（为常数）进行预测。然而，从数据来看，我们知道价格比红利的波动性更强，价格/红利比可以用来预测未来收益。简单的红利适应性学习模型能“刻画现实”吗？(哪怕是期望收益为常数的模型?）为了分析这个问题，考虑一个样本容量为 n 的关于 μ 和 σ^2 的标准适应性学习过程：

$$\hat{\mu}_t = \frac{n-1}{n}\hat{\mu}_{t-1} + \frac{\Delta\ln D_t}{n} \tag{37a}$$

$$\sigma^2 = \frac{n-1}{n^2}\ [n\hat{\sigma}_{t-1} + (\hat{\mu}_{t-1} - \Delta\ln D_t)^2] \tag{37b}$$

红利的条件预期为

$$\hat{E}_t D_{t+i} = D_t\exp\ (i\hat{\mu}_t + i\hat{\sigma}_t^2/2) \tag{38}$$

将其代入 RVF 就得到一个“学习情形中价格”的非线性函数

$$\hat{P}_t = \left\{\frac{\exp\ (\hat{\mu}_t + \hat{\sigma}_t^2/2)}{1+R-\exp\ (\hat{\mu}_t + \hat{\sigma}_t^2/2)}\right\}D_t \qquad (\hat{\mu}_t + \hat{\sigma}_t^2 < \ln\ (1+R)) \tag{39}$$

对（35）进行回归就得到μ 和 σ 的估计值，结合 R 的估计值，我们就可以模拟出 RE 下的红利序列，并计算出 P_t^{RE}。

类似地，给定初始值 $\hat{\mu}_1$ 和 $\hat{\sigma}_1$，(37)、(38）和（39）给出了“学习情形中”$\hat{P}_t$ 及 $\hat{P}_t/\hat{D}_t$ 和资产收益 $(1+\hat{R}_t) = (\hat{P}_t + D_t)/\hat{P}_{t-1}$的拟合值。在 MCS 中，(39）使用如下规则来排除无

穷多个价格的情形：如果$\hat{\mu}_t+\hat{\sigma}_t^2/2\geq\ln$（$1+R$），那么就用$\hat{\mu}_{t-1}$和$\hat{\sigma}_{t-1}^2$来代替［$\hat{\mu}_t$，$\hat{\sigma}_t^2$］。重复模拟1000次就得到$\hat{R}_{t+1}$对（$\hat{D}/\hat{P}_t$）进行回归的系数的经验分布。当然，随着$n$的增加，$\hat{P}_t$接近于序列$P_t^{RE}$，因为$\hat{\mu}_t$和$\hat{\sigma}_t^2$收敛于$\mu$和$\sigma$的总体值。

Timmermann（1993）发现，学习的存在增加了股票价格的波动性（与模拟的*RE*价格相比），即使在100年的“大”样本情形中，在学习情形中Shiller方差比被违反的可能性仍超过10%（即使模型中贴现率为常数，即使模型“在学习规则下是理性的”）。当样本容量为40时，Shiller方差比被违反的可能性高达30%～50%。对学习模型而言，当样本容量n为40～500时，收益对*D/P*的回归系数在统计上显著，所拟合的收益也不存在序列相关——这是实际数据的一个特征。

因此，所观察到的“超额波率”和收益可预测性可能不是由时变风险溢价（即预期收益不断变化）所导致的，如果允许代理人对（红利过程的）总体参数进行学习，那么可以用贴现率为常数来进行解释。

为什么学习会产生收益可预测性呢？假设代理人对红利增长的当前预测值$\hat{\mu}_t$比实际水平μ低。价格将比RE下的价格低，因此，D/P很高。当$\hat{\mu}_t$朝着μ增加时，通过（39），价格会上升（速度超过*RE*下的价格），还会生成正的未来收益，这个收益序列与今天的D/P相关。要直观地考察学习如何生成价格的“过度”波动性并不是一件很直接的事情（参见Timmermann（1993）），但可以肯定的是，$\hat{\mu}_t$，$\hat{\sigma}_t^2$中的运动幅度比μ和σ总体值的运动幅度大。另外，在RE下，红利方面的信息会导致价格的同比例调整，但在学习情形中，增长率的估计值$\hat{\mu}_t$也会增加，这会导致价格出现过度波动性。

Timmermann（1996）重复了上述方法，不过他这一次假设红利是趋势平稳序列（无论是水平值还是对数值），

$$D_{t+1}=\mu+\gamma t+\rho D_t+\varepsilon_{t+1}\quad \varepsilon_{t+1}\sim niid\ (0,\ \sigma^2) \tag{40}$$

这一点非常重要，因为Shiller（1981，1989）的实证结果表明在红利平稳的假设下方差比要被违反，而Kleiden（1986）则表明，即使RE成立，如果红利非平稳，那么方差比被违反的可能性更大。Timmermann对（40）用递归OLS得到不同样本容量下基于学习的价格$\hat{P}_t$，以便学习参数最终收敛于总体参数值（μ，ρ，γ，σ^2）。他发现在学习过程如果红利是一个平稳过程，那么即使在大样本中，MCS仍然会得到大量违反方差比的结果。收益仍然可以通过（P/D）的值来预测，且收益是自相关的。因此，这再次表明，违反真实数据中的Shiller方差比可能影响到其他理性模型（贴现率为常数）的学习过程。

Veronesi（2003）将对红利漂移率的学习与比索问题结合起来对上述方法进行了扩展。在{0，*T*}的样本数据中，红利的漂移（增长）率实际上没有变化，但在20年中增长率下降5%（20世纪30年代的衰退时期就是这种情形）的事前概率（$\lambda=0.005$）很小，也就是说，在200年时间里发生一次。一旦经济进入“衰退期”，经济恢复到正常增长率的概率也很小，$\mu=0.05$。代理人通过观察历史事实来对红利增长率进行学习。因此，飘移率和经济是否进入衰退时期都具有一定的不确定性。如果他们观察到负面的红利信息，那么处于“正常状态”的条件概率从接近1的水平下降到π（t）$=0.5$（不确定性最大的点）。当投资变得更加不确定时，他们对消息的反应更加强烈，预期收益会出现更大的波动，要求提高贴现率。假设投资者追求跨期期望效用的最大化，效用函数是关于消费的常绝对风险规避型（CARA）效用函数（$U(C)=\mathrm{e}^{-\gamma C}$），从而风险资产的需求独立于财富。RE均衡价格是$\pi$（$t$）的凸函数，当$\pi$（$t$）接近1时，这个函数非常陡峭；当$\pi(t)$接近0时，函数很平坦。因此，股票价格存在非对称效应，也就是说，股票价格对好时期的坏消息有很大的反应，对坏时期的好消息有较小的反应。假设红利遵循布朗运动（即随即游走的连续时间等价版本），因此，可以证明事后概率$\pi(t)$也遵循布朗运动，漂移率依赖于μ和θ。

模型的定性结果与股票收益和波动率的许多典型事实一致，在对模型进行校准后（例如，

每年的无风险利率为3%，隐含的相对风险规避系数$\gamma\omega$约等于3)，可以使用蒙特卡罗模拟来复制这些典型事实。相对于红利增长率为已知常数的“不存在学习，不存在比索问题”的基准情形而言，模型给出的股权溢价相对较高（0.77%～3%），这个模型就是GARCH（1，1）模型。模型有很强的持久性，对正反消息的效应不对称（如杠杆效应）。价格对红利也存在过度反应，因为π（t）会发生变化，前提条件是红利对价格存在额外的影响，这意味着弹性大于1。在“基准情形”中，价格与红利（波动率为6.5%）成比例变动，但在学习情形中，相对于红利波动率而言，股票价格的波动率会增加（从每年的6.5%增加到21%）。因此，投资者对红利增长率进行学习的比索问题可以刻画股票收益的许多典型事实，即使区制转移的概率非常小（但是，这个概率值在多年情形中相对较大且具有持续性)。

均衡模型中的学习

Timmermann（1993，1996）表明，参数的不确定性和学习能产生可预测性和过度波动性。Lewell和Shanken（2002）使用完全的理性代理人在一般均衡模型中对这种方法进行了扩展，模型中的理性代理人使用贝叶斯规则来更新参数的估计值。与Timmermann的早期结果一样，这个模型表明，即使可以用历史数据来预测收益，即使可以通过（历史数据之上的）“样本外”预测来获得利润，但从长远来看，根据历史数据的回归估计无助于事前未来收益的预测。这是因为，可预测性实际上是由学习所造成的（但这不是模型假设的一个部分）。因此，完全的理性投资者察觉不到数据中的可预测性，他们认为预期收益为常数，而不是利用可预测性来获得异常利润。这是一个很深刻的结论，因为它意味着我们对EMH所进行的绝大多数重要检验（基于历史数据对经风险调整后的利润机会是否存在进行考察）是没有用的。这一矛盾结果产生的部分原因是，重复博弈（即对历史利润进行频繁干预）的结果与投资决策的条件本质（投资决策取决于某个时点上所能观察到的红利）之间存在差异。

这个模型的结构与DeLong、Shleifer、Summers和Waldmann（1990a）非常相似，读者在进行深入学习之前可以先到第19章去看看这个模型。我们学习模型中的个体生活两个时期，代际之间存在重叠，投资者的单期绝对风险规避（CRRA）效用函数是期末（第二期）财富的函数，$U(\omega) = -E^{-2\gamma\omega}$，最优资产份额$\alpha_t^*$为

$$\alpha_t^* = \frac{1}{2\gamma}\left[\mathrm{Var}_t^s(p_{t+1}+d_{t+1})\right]^{-1}\left[E_t^s(p_{t+1}+d_{t+1}) - (1+r)p_t\right] \tag{41}$$

这里的$\gamma>0$是风险规避系数，上标s表示投资者的主观分布，投资者使用他们认为重要的信息，如风险资产的价格P_t（不是价格的对数值）、红利水平d_t。无风险利率为γ，无风险资产具有完全的供给弹性，每种风险资产的供给（标准化）为1。红利d_t的真实分布（对投资者来说是未知的）关于时间是$niid$，均值为δ，方差为σ^2

$$d_t \sim N^{true}(\delta, \sigma^2) \tag{42}$$

因此，当$t\to\infty$时，所估计的风险也接近0。在第一期中，个体对外生的财富进行配置决策。我们假设只有一种风险资产（因为我们不讨论CAPM的含义和资产收益的横截面行为，这些内容可以参见Lewellen和Shanken（2002)）。因为风险资产的供给为1，所以用下一期的资产价格所表示的均衡价格为

$$p_t = \frac{1}{1+r}\left[E_t^s(p_{t+1}+d_{t+1}) - 2\gamma\mathrm{Var}_t^s(p_{t+1}+d_{t+1})\right] \tag{43}$$

在完全信息（没有不确定性）条件下，均衡价格为（$p_t = p$为常数）为

$$p = \frac{\delta}{r} - \frac{2\gamma\sigma^2}{r} \tag{44}$$

关于参数不确定性的最简单假设是，投资者对δ没有先念看法（即“扩散式”先念)，因此他们对δ的估计只依赖于历史数据（我们也假设σ^2已知）。使用贝叶斯法则，δ在t时刻的后验分布为$N(\bar{d}, (1/t)\sigma^2)$，主观“预测”分布为

$$d_{t+1} \sim N^s\left(\bar{d}_t,\ \left(\frac{t+1}{t}\right)\sigma^2\right) \tag{45}$$

投资者对 δ 的最优预测就是红利的平均值 $\bar{d}_t$（使用直到 t 时刻的数据）。d_t 均值的不确定性在预测方差中有所反映。理性投资者都是贝叶斯主义者，但也都具有前瞻性，因此求解价格方程就可以得到当前价格的闭型解

$$p_t = \frac{\bar{d}_t}{r} - 2\gamma f(t)\ \sigma^2 \tag{46}$$

$f(t)$ 是关于时间的定性函数，当 $t \to \infty$ 时，$f(t) \to 1/r$（因为 δ 和 σ^2 都为常数）。

我们现在正在对贝叶斯学习基础上的主观收益分布和真实收益分布进行分析。信仰贝叶斯主义的投资者的行为决定了价格的实际变化

$$p_{t+1} - p_t = \frac{1}{r}\ (\bar{d}_{t+1} - \bar{d}_t)\ + 2\gamma\ [f(t)\ - f(t+1)] \tag{47}$$

忽略这些确定性项，投资者认为价格服从一个鞅过程

$$E_t^s\ (p_{t+1} - p_t)\ = 0 \tag{48}$$

因为投资者无法预测（主观的）预期红利的变化。然而，当投资者根据新获得的数据来更新关于 $\bar{d}_t$ 的看法时，真实价格具有时变性。

如果我们假设红利的真实平均值为 $\delta = 0.05$（每年为 5%），$\sigma^2 = 0.10$（每年 10%），那么可以用（47）来生成学习情形中的价格序列，其中 $\bar{d}_t$ 根据正态分布的随机抽取来确定。在完全信息情形中均衡价格 $p_t = 1$（对所有的 t 都成立），而在不完全信息情形中，价格围绕“1”进行长期波动。但是，当 $t \to \infty$ 时，价格接近于基础价值 1（因为在贝叶斯学习中，常参数收敛于参数的真实值）。这些拟合价格“看上去”不像随机游走，因为它们依赖于 $\bar{d}_t = d_0 + d_1 + \cdots + d_t$，（其中 $d_t \sim N(\delta,\ \sigma^2)$），这些拟合价格具有长期波动特征，价格最终会收敛到真实的均值 δ。因此，从数据中可以观察到过度波动性。然而，贝叶斯投资者无法完全预测价格的变化。假设价格的实现方式是给定的（在学习情形中），可以证明，“绝对”收益（absolute return）是序列负相关的，$\mathrm{Cov}(R_t,\ R_{t+1}) < 0$，其中 $R_{t+1} = d_{t+1} + p_{t+1} - p_t$。另外，今天的高红利可以预测下一期的低收益（即 $\mathrm{Cov}(d_t,\ R_{t+1}) < 0$），收益在长期中存在均值反转（因为价格最终等于基础价值）。因此，即使所有投资者都是理性的且观察不到价格中的可预测性，计量经济学家还是可以根据数据来“揭示”这些特征。

我们现在转到如下问题的讨论：贝叶斯投资者始终认为预期利润为 0，为什么我们使用基于历史数据的“样本外”检验还是可以发现一些获利的交易规则？假设交易法则为，如果 $\bar{d}_t < K$（$> K$），那么就持有风险（无风险）资产，其中 K 是过滤法则中的任意常数。对历史数据进行重复抽样（即频繁观察），如果 $\bar{d}_t < K$，那么 $\bar{d}_t$ 也小于真实的平均值 δ，因此，平均来看，已经实现的红利和下一期的价格都会更高。但贝叶斯投资者都是根据从历史红利过程中抽取的数据为基础来进行决策的。投资者对 δ 的看法都是以 $\bar{d}_t$ 为中心，$\bar{d}_t$ 大于 δ 和小于 δ 的可能性相同。重复抽样对贝叶斯主义者而言并不重要。

上述模型是 Lewellen 和 Shanken 模型最简单的一个版本，他们也考虑了信息先念和已知的状态变化，考察了对收益的横截面方差进行检验模型的含义（即使投资位于均值方差有效前沿上，CAPM 也会遭到拒绝）。因此，考虑理性投资者和参数不确定性的模型与典型事实是一致的，尽管这并不能证明学习是造成这些结果的原因。因此，参数的不确定性使得我们对市场有效性检验的解释变得相当危险（hazardous）。基于“真实数据”对 EMH 更为严格的检验不仅需要均衡收益模型，还需要一个能模拟贝叶斯更新法则和合理“先念”的模型——这是一项令人生畏的工作。

尽管我们在这里不对这些东西进行纠缠，还是要注意的是，当我们根据跨期学习模型来导出资产需求时，参数不确定性所扮演的角色与其他状态变量（如随机劳动收入）类似，这会导致资产的跨期对冲需求（在 Lewellen - Shanken 模型的资产需求函数中不会出现这种对冲需

求）。我们在后面的跨期模型一章中将简要地讨论这个复杂问题。

自我参考学习（self - referential learning）

在上述模型中，红利服从外生的时变过程，代理人对红利过程的估计最终会收敛于它的真实值，因此价格（通过 RVF）最终也会收敛于它的真实值。现在假设红利和价格之间存在反馈机制——学习过程的收敛性不再得到保证，也可能存在多重理性预期均衡。自我参考学习模型可用来解决这个问题（Marcet 和 Sargent（1989））。

我们可以认为代理人有一组参数的递归估计值 $\hat{\rho}_t$，因此数据由函数 $T(\hat{\rho}_t)$ 来生成。如果 z_t 是代理人所使用的可观察状态变量，那么 z_t 动态行为的一般表示是

$$z_t = T(\rho_t)\ z_{t-1} + V(\rho_t)\ \mu_t \tag{49}$$

其中，μ_t 是影响 z_t 的一个“冲击”，影响程度依赖于代理人的参数估计 $V(\rho_t)$。理性预期解是 $T(\rho)=\rho$ 中的不动点 ρ。向理性预期收敛的稳定性依赖于如下微分方程的稳定性

$$\frac{\partial \rho_t}{\partial t} = T(\rho_t) - \rho_t \tag{50}$$

在不动点附近进行线性化就可以得到局部稳定性

$$\frac{\partial \rho_t}{\partial t} = T_\rho'(\rho)\ (\rho_t - \rho) \tag{51}$$

局部稳定性取决于特征值 $T_\rho'(\rho)$ 的实部。从直观来看，待估参数 ρ_t 必须接近于均衡参数，因此，代理人的先念看法必须排除那些不使参数朝均衡值方向运动的估计值。显然，自我参考学习对代理人的要求比简单（OLS）自适应学习复杂得多，简单自适应过程忽视了反馈机制。

在股票价格情形中，自我参考学习的稳定性对由 RVF 所给出股票价格的简约形式施加了限制条件。为了对其进行简略概述，假设我们对由红利和价格所组成 Var 进行估计（Timmermann（1994）），并使用预测值

$$\hat{D}_{t+1} = \hat{\phi}_{1t} D_t + \hat{\lambda}_{1t} P_t \tag{52a}$$

$$\hat{P}_{t+1} = \hat{c}_{1t} P_t + \hat{c}_{3t} D_t \tag{52b}$$

假定潜在的真实红利过程中的参数为常数，

$$D_t = \phi_1 D_{t-1} + \lambda_1 P_{t-1} + u_t \tag{52c}$$

欧拉方程为

$$P_t = \theta E_t\ [P_{t+1} + D_{t+1}] \qquad (0 < \theta < 1) \tag{53}$$

将（52a）和（52b）代入（53）并使用（52c）来消掉 D_t 就得到

$$\begin{aligned} P_t = \frac{\theta}{1 - \theta\hat{c}_{1t} - \theta\hat{\lambda}_{1t}} \{ & [\lambda_1 (\hat{c}_{3t} + \hat{\phi}_{1t})]\ P_{t-1} \\ & + [\phi_1 (\hat{c}_{3t} + \hat{\phi}_{1t})]\ D_{t-1} + (\hat{c}_{3t} + \hat{\phi}_{1t})\ u_t \} \end{aligned} \tag{54}$$

注意，（54）包含由真实（常数）系数（ϕ_1，λ_1）和根据（52a）和（52b）所得到的估计值所组成的项。RE 价格依赖于（P_{t-1}，D_{t-1}），但是，因为代理人对参数的了解不全面，随着参数的依次更新，P_t 依赖 P 与 D 和的所有滞后值。在学习情形中（正如 RE 所反对的那样），历史对价格的动态行为很重要。状态“向量”为 $z_t =$（P_t，D_t），在双变量 Var 中可以使用 OLS 来估计 $\rho_t =$（$\hat{\phi}_{1t}$，$\hat{\lambda}_{1t}$，$\hat{c}_{1t}$，$\hat{c}_{3t}$）。理性预期意味着，在（52a）和（52b）所包含的信息给定的条件下，代理人的预测必须等于用（52c）和（54）所表示的真实数据生成过程中的系数。对（54）和（52b）进行比较，价格方程的不动点就是如下方程组的解：

$$T_{11}(\rho) = \frac{\theta\lambda_1 (c_3 + \phi_1)}{1 - \theta c_1 - \theta\lambda_1} = c_1 \tag{55a}$$

$$T_{12}(\rho) = \frac{\theta\phi_1 (c_3 + \phi_1)}{1 - \theta c_1 - \theta\lambda_1} = c_3 \tag{55b}$$

对红利方程而言，我们（简单地）有

$$\phi_{1t} = \phi_1 \quad 且 \quad \lambda_{1t} = \lambda_1 \tag{55c}$$

这些参数不受学习的影响并收敛于它们的真实值。方程（55）的不动点包含了通常的RE非线性交叉方程限制。朝RE的收敛性取决于价格方程在（55a）和（55b）不动点处导数的特征值。Timmermann（1994）使用英国的股票价格指数数据（1926—1986）来对（52a）和（52b）所表示的Var进行了估计，计算出了进行线性化后的微分方程的特征值。他发现，就（52）中的系统（水平值形式）而言，特征值表明学习过程不收敛，但是，当我们对一阶差分形式的系统进行估计时，有一个不动点是稳定的，所以代理人可以通过“学习”来逼近RE均衡。显然，这两组相互矛盾的结果说明水平值形式的方程组中变量不是平稳的。

在后来的文章（Timmermann（1996））中，他用美国S&P500指数数据对Var进行递归估计，并用（54）来构造一个基于自我参考学习的价格P_t（使用MCS）。他发现，与外生的自适应学习（即仅对红利方程进行递归OLS估计）相比，在有限样本数据中，自我参考学习中的价格序列的波动性比RE价格序列小。假设自我参考学习模型不得不接近于RE模型，如果收敛最终会发生，那么也许这并令人感到奇怪。因此，“外生”的OLS学习模型比自我参考学习模型更适合对过度波动性这一典型事实进行“拟合”。

状态变化与学习

现在我们用Campbell－Shiller的VRA方法来对美国股票价格进行研究，这种方法涉及状态（区制）变化和红利预测与未来预期收益之间的差异，这些都取决于投资者认为自己所处的状态。超额波动性的产生原因、未来红利和预期收益在股票价格运动解释中的相对重要性都可以用来与“单一状态”下的固定系数Var进行比较。

我们已经注意到，（对数）红利是否平稳不能一概而论，所观察到的数据序列中可能存在状态变化（如“高”和“低”红利增长时期）。因此，代理人必须明白他们处于哪一种状态中，这有可能对他们对未来红利增长的预测产生影响。这是Evans（1998）的思想，他对Campbell－Shiller线性化后的RVF进行了考察，假设由“基础变量”（$r_t - \Delta d_t$）变化的预测值和红利价格比δ_t变化的预测值所组成的Var有Hamilton（1989）的两状态区制变化表示。就1871—1991年期间美国的年度数据而言，由区制转换Var所得到的预测值与由标准（参数为常数）Var所得到的结果存在很大的差异。简单地说，区制转换Var允许基础变量具有更大的波动性（因为状态转移），因此，这个模型能更好地解释（对数）红利价格比δ_t所等观察到的变化，其中

$$\delta_t = \tilde{E}_t \sum_{i=1}^{\infty} \rho^{i-1} \left(r_{t+i} - \Delta d_{t+i} \right) + 常数 \tag{56}$$

r_t为利率，ρ为常数，期望$\tilde{E}_t$取决于预期的未来状态。个体的预测误差取决于RE预测误差和对未来状态进行预期的误差（后者在小样本中不为0）。因为状态转移导致了投资者的预测误差与红利价格比δ_t之间存在小样本相关性，因此，在这个框架下，即使RVF成立，也有可能出现的情形是Shiller方差边界被违反。区制转移Var模型意味着δ_t变化的60%源自基础变量（$r - \Delta d$）的变化（在标准Var中为35%），只有14%的变化源自预期收益的变化。（模型的扩展允许代理人对他们当前所处的状态进行学习，但这不会对实证结果产生太大的影响。）因此，作为股票价格变动的一种解释，（与状态转移的模型相比较）在状态转移模型中，“现金流”相对而言要重要一些，而时变预期收益的重要性要小一些。这说明在使用不同模型时要得到与“典型”事实一致结论非常困难。在区制转移模型中，事后股票收益在长期中具有可预测性。

不是经常性的结构突变

在Barsky和DeLong（1993）和包含学习的早期模型中，总体参数（如红利增长率）为常

数。在 Evans（1998）中，红利在两个状态之间进行转换，投资者面临的唯一问题就是识别当前未知状态的滤波问题。例如，当我们考虑经济周期所处的阶段时，后者似乎是合理的，但是，很少出现如下情形（如 1929—1932，二次世界大战）：红利过程经过一次“突变”后不再次发生，突变预示着一种新状态的产生。Timmermann（2001）认为，在这些“突变”之后，投资者将面临更大的不确定性，因为历史数据几乎没有价值，这些突变很有可能导致对预期产生较大的修正。如果投资者对突变具有完全信息（如红利的增长率更高），那么这会导致收益过程在突变时刻存在一个异常值。因此，Timmermann 认为，代理人对突变的学习是非常缓慢的（即不完全信息），这对收益分布而言具有重要意义（如偏斜、尖峰、自相关、波动聚集效应）。

Timmermann 的方法是建立一个模型，模型中红利的低频率突变受区制转移 Markov 过程所主导，后者可以扩展到非重复发生的状态集合上去。然后，模型可以在完全信息、贝叶斯学习（以已知的状态为条件）和“滤波模型”（通过贝叶斯规则来更新状态发生的概率）等条件下来求解股票价格。当我们使用相对风险规避系数为常数的幂效用函数时，股票价格的欧拉方程为（$D_{t+1}=D_{t+1}$）

$$P_t=\delta E_t\left[\left(P_{t+1}+D_{t+1}\right)\left(D_{t+1}/D_t\right)^{-\gamma}\right] \tag{57}$$

实际红利由下式给出

$$\ln D_{t+1}=\ln D_t+\mu_{t+1}+\sigma_{t+1}\varepsilon_{t+1} \tag{58}$$

μ_{t+1}和σ_{t+1}在突变点会发生变化，$\varepsilon_{t+1}\sim N(0,1)$，因此红利增长率为 $g_{t+1}=\exp(\mu_{t+1}+\sigma_{t+1}^2/2)-1$。

在突变发生后，新的平均增长率 t_{t+1} 可以从一个一致密度函数（uniform density function）$G(\cdot)$中进行抽取，密度函数的上下限为 $[\underline{g},\bar{g}]$。在模拟结果中，对 1871—1999 年的数据而言，（实际）红利增长率的最小值和最大值分别为 −4% 和 6%（每年），平均增长率为 1%（每年），每月的方差为 $\sigma=1.5\%$（每年 5.2%）。实际贴现率 $\delta=7.5\%$（每年），状态之间的转移概率为 $\pi=0.997$，这意味着，对红利过程而言，每 30 年会发生一次漂移（在 1871—1999 年的模拟时期中，转移的平均次数为 4 ~5 次）。

我们知道，月度股票收益（S&P500）具有高波动性、偏斜、肥尾、一阶序列相关和波动率聚集（ARCH）效应等特征。在不存在突变（$\pi=1$）的模拟模型中，这些特征都不会出现（因为 P_t 受平稳的参数固定的随机过程所控制）。当突变存在且信息完全时（$\gamma=0$），由突变所造成的收益异常值会导致波动率、偏斜程度和尖峰程度等都有所增加，但波动聚集效应不会出现。

在贝叶斯学习中，模型所得到的平均波动率为 3.3%，接近于样本估计值 4.1%，偏斜、尖峰和波动聚集效应（对 $\gamma<0.5$ 的情形而言）都与数据中所发现的结果相似，但收益不存在序列相关性。Timmermann 还用模拟数据和 GMM 估计法中的工具变量 $z=$（常数，红利收益，滞后收益）来对欧拉方程（49）过度识别限制条件进行了检验。他发现，当“突变”和学习存在时，在 5% 的显著性水平下对 Hansen 的 J - 检验拒绝频率达到 10%。

我们再次发现，与收益有关的大量“典型”事实与我们的模型一致，模型中存在红利过程的突变，代理人不得不缓慢地更新对未来红利的预测。

17.5 小结

* 从数学角度来看，理性泡沫的产生是因为，当没有横截性条件约束时，由欧拉方程所得到的股票价格的解等于基础价值加上一个“泡沫”项，最后一项服从鞅过程。

* 当泡沫存在时，股票收益不可预测，因此 RE 的正交检验不能用来对理性泡沫进行检验。

* 在早期文献中，泡沫相对于基础变量（如红利）而言是外生的。泡沫的起源无法解释，这些模型只是给出了泡沫的时间轨迹。

* 标准单位根检验和协整检验可用来对连续的爆发性泡沫进行检验，但不可能用来对周期崩溃性泡沫进行检验。

* 内在泡沫以非线性的定性方式依赖于经济中的基础变量（如红利），内在泡沫仍然满足欧拉方程。内在泡沫证据的部分来源是对泡沫解直接进行估计，这里的价格依赖于红利的非线性函数。内在泡沫的证据无法得出确切的结论。

* 如果价格由 RVF 来确定，但代理人递归地更新他们对红利过程真实（固定）参数的估计值，那么“学习情形中的价格”比“纯”理性预期情形中的价格波动性更高，即使预期收益为常数，结论也不会发生变化。随着参数的递归更新，模拟收益自相关，红利价格比可以用来预测收益（这是对真实数据的模拟）。

* 在很少发生但“结构性突变”比较剧烈的模型中，代理人在非重复发生的状态发生变化后必须对更新后的参数进学习和了解，这种模型能模拟实际收益数据中所发现的收益的一些典型事实：高波动性、偏斜、尖峰、自相关和波动聚集效应。

* 如果代理人是理性的，使用贝叶斯法则来更新参数，那么均衡模型仍然能复制上述典型事实，特别地，人们能从历史数据中找到可获利的交易法则即使投资者在事前无法发现这些交易法则。

* 在“自我参考学习”的情形中，价格是否朝“纯”RE 均衡进行收敛依赖于学习模型中参数的大小。美国（S&P500）数据所给出的股票价格波动率不会比纯 *RE* 模型中的波动率高，因此，自我参考学习模型似乎不能解释现实股票价格中高波动性这一“典型事实”（贴现率为常数）。

* 在考虑“基础变量在不同状态之间进行转移”的模型中，代理人面临“过滤”问题（代理人为了预测未来红利必须确定处于某种状态的概率）。在 Compbell - Shiller 的 Var 方法中，状态转移模型所得到的预测结果（如对红利增长进行预测）与“固定系数”Var 模型所得到的结果不同，因此，在解释股票价格运动时就会得出红利增长和时时变预期收益相对重要性的不同结论。

第18章

行为金融与异像

目的

- 说明行为金融与标准有效市场方法有什么不同。
- 说明风险规避、有限投资期限和系统风险如何导致错误定价。
- 说明噪声交易者如何在市场中存活。
- 对来自个体和群体行为的心理学实验证据进行分析。
- 概述在股票市场中所发现的主要异像，说明行为金融如何“解释”潜在的获利机会。

18.1 主要思想

在一个特殊的市场上，只有当它与某些基于基本面因素的选择模型相结合时才能公然宣称市场是无效的。有效性是正确模型和直接排除实际价格与基础价值（由模型决定的价格）之间差异的联合假设。模型不能对现实进行完美刻画，因此总是存在一个“残差”——整个数据集中有一些不能被模型解释的因素。如果这些“残差”拥有系统性模式，那么我们就称其为异像或谜。行为金融要么通过某些非标准行为（如非理性或非标准偏好）要么借助理性投资者设法使公平价格（基础价值）与实际价格相等的事实来对这些异像进行解释。“噪声交易者”（noise traders）是相对于非基本面交易者（non - fundamentals traders）而言的一个术语。

在许多行为金融模型中，投资者在预算约束和其他约束（如不允许卖空）条件下仍然追求期望效用的最大化，但所选择的效用函数和投资者形成预期的方式都不是标准的。在其他类型的行为金融模型中，存在两个不同的群体：理性的基本面交易者和“非理性的”噪声交易者。

假设理性经验事实是“合理的”，在“解释”经验事实方面，理性基础变量方法（存在交易成本）和各种行为方法之间到底存在多大的差距，这是一个争论很大的问题。例如，当我们谈到一个足球运动员（如大卫·贝克汉姆）的表现时，“基本面支持者”会说：“因为足球场不是足够的平整，所以他得分的情况没有用个体利润最大化模型所预测的好”。行为模型的支持者会说：“他得不到那么多分是因为，当别人进球比他多时，他丧失了自尊心。”（也就是说，他的效用函数除了目标得分外还多了一个变量，他属于损失规避型（loss - aversion））。行为金融模型的支持者也可能会说：“追求得分最大化的目标函数和个体利润都是正确的，但他的对手不是理性的（例如妒忌他的名誉和地位），因此，他们扮演的是噪声运动员的角色，他们使用不正当手段来伤害他。”所以，与完全理性的情况（即所有队员都按规矩踢球）相比，他的得分情况要差一些。这种方法认为，足球比赛的结果取决于“理性”（跨期利润最大化）

运动员和非理性“噪声”运动员之间的交互影响，也取决于前者是否模仿后者的行为。（大家也许已经注意到，足球场上存在一定数量的“针锋相对”行为，当一方开始犯规时，另一方得到观众的喋喋声并公然在裁判面前疯狂犯规——例如，在 1998 年世界杯上英格兰对阿根廷的比赛中贝克汉姆被发下场的时刻。）

由于模型可能存在误差，反对 EMH 的大量证据都是基于所谓的“异像”，而不是基于基本面来对模型的复杂统计检验（如对 CAPM 的回归检验、Shiller 波动率检验等）进行评价。为了使模型的误差达到最小，文献的一个分支是考察很接近的可以相互替代的资产（如孪生股票（twin shares），美国存托凭证（ADRs），封闭式基金等）之间的定价差异，而这个差异可以通过套利交易来进行消除。异像文献的另外一个分支是寻找对风险和交易成本进行修正后的能赚钱的交易规则——这需要风险的统计测度指标，而这些指标存在很大的争议，因此，风险的测度仍然涉及一些“模型误差”。行为模型要回答的两个关键问题是：

* 为什么市场会持续出现错误定价？

* 噪声交易者在市场中如何生存？

在行为模型中没有“免费午餐”（即经过风险和交易成本调整后的超额利润）。但是，这并不意味着“价格就是正确的”（即价格等于预期现金流的贴现值，这种预期服从正确的分布，贴现率以可接受的偏好假设为基础）。从本质来看，“价格是正确的”是 EMH 的一种简略说法。但在某些行为模型里，因为理性交易者缺乏消除噪声交易者导致的错误定价的方法，所以价格不等于基础价值。

注意，EMH 不要求市场中所有参与者都是“有效的”且信息灵通。市场中可能存在一些非理性交易者或“噪声”交易者，他们的报价不等于基础价值。EMH 需要足够多的“精明”交易者，他们能认识到 P_t 最终会等于基础价值 V_t。因此，如果一些非理性交易者的报价 $P_t < V_t$，精明交易者会马上介入并从非理性交易者手中购买股票，这就使得 P_t 很快地向 V_t 运动。

举一个简单的例子，假设所有足球俱乐部都通过发行股票来进行融资（读者如果对其他运动感兴趣，可以用自己喜欢的球队或者运动员来进行替换）。假设噪声交易者在开始的时候以公平价格 25 英镑购买了皇家马德里俱乐部的股票。现在假设 n 小时以后噪声交易者作为一个群体在酒吧里喝多了酒，他们对皇家马德里俱乐部的管理层存在非理性的过分悲观情绪，他们最终以价格 $P=20$ 英镑卖出了自己的股票，该价格低于基础价值 $V=25$ 英镑，所以他们亏了本。信息完全的理性投资者（如维多利亚 · 贝克汉姆女士）就会介入并购买皇家马德里的股票，与此同时，为了进行套期保值，她卖空（正确定价）可以近似替代皇家马德里股票的资产，这种资产与皇家马德里有相同的未来现金流。假设这种替代资产是巴塞罗那股票（历史上的竞争者）。如果足球俱乐部的市场价格普遍下降（如电视广告收益损失），那么当错误定价得到修正时，“多 - 空”的理性交易者仍然能获得利润。市场价格普遍下降意味着皇家马德里的股票价格也在下降，比如说 1 英镑的损失，但可以从巴塞罗那股票的空方头寸上获得 1 英镑的收益。但是，当皇家马德里的错误定价得到修正时，“多 - 空”交易者可以获得每股 5 英镑的收益。这是因为，如果理性交易者进入市场并以 20 英镑的价格买入股票，那么他们会将股票价格推向 25 英镑的公平价值，他们在这个过程中可以获得利润。（利润以噪声交易者的损失为代价。）噪声交易者最初以 25 英镑的价格买入股票，然后以低于 25 英镑的价格卖出（给理性交易者），他们最终会被理性交易者“挤出市场”（达尔文优胜劣汰的生存法则）。如果有足够多的理性投资者和足够多的资金愿意进入市场，那么皇家马德里的价格会很快回到它的基础价值。然而，上述情景因为下述原因可能不会发生。

基础风险

卖空替代证券可以避免理性交易者遭受（足球）产业（系统）风险，但特殊风险（如效力于皇家马德里俱乐部的球员大卫 · 贝克汉姆的腿受伤的新闻）仍然存在。因此，“套利”并不是没有风险。

噪声交易风险和有限投资期限

即使巴塞罗那是皇家马德里的完美替代，噪声交易者还是有可能对皇家马德里的管理层更加悲观，由此可能导致皇家马德里的价格进一步下降（低于20英镑）。这就是理性交易者所面临的噪声交易风险，如果理性交易者的期限较长，价格最终收敛于基础价值，那么他就不会担心噪声交易者风险，他们要做的就是为套利利润机会的到来等待足够长的时间并避免短期损失。但是，职业投资组合管理人员一般都只有相对较短的期限来管理他人的资金，因此，“脑力和资本是分离的”（Shleifer 和 Vishny（1997）），从而套利机会不会消失。

卖空

绝大多数养老基金和共同基金的管理人员是不允许进行卖空操作的，对冲基金却可以进行卖空操作，但必须从经纪人那里获得丰富的股票（巴塞罗那）供给。一旦错误定价得到修正，对冲基金就必须借到（巴塞罗那）股票。如果不能做到这一点，那么他们就不得不在市场中以不利的价格来回购巴塞罗那股票——这就是所谓的“买入”。另外，如果涉及外国股票（如国际米兰），那么卖空是很难做到的，经纪人佣金和竞买竞卖差价在进行套利计算时都必须进行考虑。（这与米兰足球“市场”不是皇家马德里股票市场风险的近似替代一样。）

模型风险

理性交易者不能确信皇家马德里20英镑的股票价格是否被低估，因为他们所估计的基础价值为25英镑（如使用Gordon增长模型）可能是错误的。这种“模型风险”可能会限制理性套利者的头寸。

选美竞赛

如果理性投资者的期限有限，那么他们将关心（未来）N时刻的价格。但是，如果他们将预期E_tP_{t+N}建立在$t+N$期之后红利的基础上，那么我们就回到无限期界的理性投资者的情形（参见第4章）。然而，如果模型中代理人不同质，那么在代理人认为世界不是由理性投资者所主宰的条件下，$t+N$时刻的价格将部分地依赖于理性投资者如何看待其他理性投资者对P_{t+N}的看法（即Keynes选美竞赛（beauty contest））。如果理性投资者知道其他理性投资者使用的是均衡收益模型，那么上述观点也是适用的。我们在这里拒绝EMH假设：所有投资者立即知道正确模型。在这种情形中，理性投资者也许会认为，真实价格是理性定价和非理性交易者（如技术交易者，诸如图表主义者、使用K线图或神经网络图的交易者，参见Cuthbertson和Nitzsche（2001a））效应的加权平均。因此，价格不会总是等于基础价值。

学习成本

如果在发掘真实模型或获得所有信息的过程中存在学习成本，那么这会限制那些持有即期头寸的理性交易者数量，从而能消除皇家马德里股票中的错误定价。可能的情形是，有足够多的套利者，总量资金也是足够的充裕，因此即使在长期中风险获利机会也会因为套利而消失。然而，如果我们意识到任何单个套利交易者都不可能知道证券的基础价值，也不可能知道所观察的价格在什么时候偏离基础价值，那么后一观点的力量会很弱。套利者作为一个群体也有可能对基础价值的看法不一致（即他们具有不同的预期），即使从长期来看这也会增加他们察觉获利机会的不确定性。因此，精明交易者很难识别市场中的错误定价。如果资金受到限制（也就是说，比套利者低估证券的需求弹性低），或者期限是有限的，那么可能的情形是，能获利的风险套利机会在市场中能持续一段时间。

人们意识到“信息成本”可能很大，而边际成本随着交易范围和数量的增加而增加，这会对某些市场中的套利活动施加限制条件。例如，我们考虑一个极端情形，如果信息成本是如

此的高以至于交易者所关注的对象要么只是债券要么只是股票（即市场完全分割），那么债券和股票预期收益之间的差异不会因为套利而立即消失。

充分必要条件

对金融市场的大量分析都依赖于套利原理（参见 Shleifer 和 Summers （1990））。套利者（或精明交易者或理性投机者）对市场进行连续观察，迅速消除实际价格与基础价值之间的差异，因此也就立即消除了任何获利机会。如果证券具有完全替代品，那么套利是无风险的。无风险套利使得相对价格相等。然而，如果资产没有很接近的替代品，那么套利行为是有风险的，套利交易无法职称股票（或债券）的绝对价格水平。

只有当精明交易者的期限有限时，有风险的套利交易头寸才可能发生。精明交易者认为，价格最终要向基础价值方向运动，因此，从长期来看，交易者可以获得利润。然而，如果套利者不得不借入现金或证券（为了卖空）来进行交易、定期支付相关费用或向管理层定期（如每月或每季度）报告“账面”利润的头寸，那么即使绝大多数交易都是精明交易者的交易，无限期界假设都不是很适用。

首先考虑一个例子：不存在近似替代证券，套利者容易暴露于基础风险之中。错误定价具有一定的持续性，如果（参见 Barberis 和 Thaler （2003））

（ⅰ）套利者属于风险规避类型；

（ⅱ）基础风险具有系统性。

条件（ⅰ）保证了没有一个套利交易者愿意持有大的头寸来消除定价误差，条件（ⅱ）保证了大量投资者的存在也无法消除定价误差，这些投资者所持有的证券头寸都很小。

如果证券存在完全替代品（即任何状态下的现金流相等），两种证券的价格却不同，那么不存在基础风险或模型风险，只存在噪声交易者风险。套利者能够完全肯定定价存在误差。要使定价误差持续存在，我们要求：

（ⅰ）套利者属于风险规避类型，持有短期头寸；

（ⅱ）噪声交易者的风险是系统风险

在（ⅰ）中，我们需要一个“短期期限”来避免出现如下情形：单个（富有的！）套利者有足够的时间来等待定价误差得到修正。条件（ⅱ）使得大量小投资者无法分散错误定价的风险。

18.2　信念与偏好

我们在上面解释了为什么风险套利会受到限制、无法保证股票的实际价格等于基础价值。我们现在来讨论个体信念和偏好方面的经验证据（参见 Barberis 和 Thaler （2003）、Shleifer 和 Summers （1990）、Shiller （1989） 的综述）。

心理学试验表明，个体会犯系统性（即非随机的）错误。试验对象对新信息（消息）存在过度反应，他们喜欢从历史价格趋势中挖掘信息。过于自信使得他们承担了较大的风险。

行为心理学的试验证据表明，人们在对未来进行预测时总是不理性。贝叶斯方法通常是最优的，个体在进行最优预测时会将先念信息和最新信息进行加权。但是，在实践中，个体似乎把当前数据看得太重，对他们的预测能力过于自信。

当代理人不知道真实的数据生成过程时，他们通常认为小样本能反应“总体特征”（即“小样本规律”，或“小样本容量问题”）。例如，如果大卫·贝克汉姆连续三场比赛都得了分，那么他处于“状态最好的时期”，人们会错误地增加“他在下一场比赛中会进球”的主观概率。但是，这种估计根本不能代表他在以前赛季中的平均得分率，事实上，他的得分模式可能不存在序列相关性。即使他们知道真实的数据生成过程（如投掷硬币），如果连续5次出现的

都是头面朝上，那么他们也会认为下一次出现反面的概率大于 1/2，这就是所谓的赌徒谬误（gambler's fallacy）。

人们经常过多地强调初始状况。当问及联合国中非洲国家加入的比例为多大时，人们的回答会受到初始给定的随机数目 x（$0<x<100$）的影响。当所提示的参考值为 10% 时，他们的答案是 25%；当所提示的参考值 $x=60\%$ 时，他们的答案是 45%（平均）。人们也可能存在记忆偏差，更近或更为突出的事件对人们就事件结果的看法有很大的影响。例如，经历过汽车被偷事件的人会过高地估计城市汽车被盗事件的概率（与没有经历过被盗事件的人相比）。

Ellsberg 悖论

有大量的实验表明，个体在各种赌博之间进行决策时，即使知道结果的真实概率分布，也经常不遵循预期效用理论的基本公理。在描述在不确定性条件下的理性选择时，即使考虑主观概率而不是用客观概率也不能拯救预期效用公理（Savage（1964））。Ellsberg（1961）的实验表明，个体不会按照预期效用公理来进行理性行事。假设有两口缸，每个缸里都装有红球和蓝球。第一口缸里有 100 个球，但不知道红球和蓝球的具体个数；第二口缸里有 50 个红球和 50 个蓝球，因此红球和蓝球随机抽到的概率都为 1/2。实验对象明白上述事实，并被要求从中进行选择，赌博的偿付是 100 美元。第一个选择是：

（a）从第一口缸中进行随机抽取，如果为红色，就得到 100 美元，否则就得到 0 美元；

（b）从第二口缸中进行随机抽取，如果是红色，就得到 100 美元，否则就得到 0 美元；

实验对象倾向于选择（b），（b）中的概率是已知的，篮球和红球出现的概率都等于 1/2。选（b）而不选择（a）意味着他们认为在第一口缸中红球出现的概率小于 1/2。

第二个选择是：

（c）从第一口缸中进行随机抽取，如果为蓝色，就得到 100 美元，否则就得到 0 美元；

（d）从第二口缸中进行随机抽取，如果为蓝色，就得到 100 美元，否则就得到 0 美元；

试验对象倾向于选择（d）而不是（c）。这意味着他们认为第一口缸中蓝球出现的概率小于 1/2，这又反过来意味着红球出现的概率大于 1/2。这与前面选择（b）而不是（a）的情形形成对比，这就是 Ellsberg 悖论。这个实验也表明，人们尤其不喜欢结果的概率分布不明确的情形。这就是通常所指的“模糊规避”（ambiguity aversion），在许多试验中都可以观察到这一现象，尤其是参与者在对各种赌博之间进行选择之前意识到自己对相关知识缺乏的情形（例如，一个对运动不感兴趣的人对运动会的比赛结果打赌（Fox 和 Tversky（1995）））。

心理账户

我们应怎样概括人们对股票市场中特殊结果所获得效用的评价呢？这里有几种可能性。假设你一开始就拥有金融财富 1000 美元，你将参加有两种可能结果的赌博，如果出现第一种结果，你将获得 200 美元，如果出现第二种结果，你将损失 50 美元。标准的终生效用函数方法意味着，$U=U(1000)+U(1200)+U(1150)$，$U(\cdot)$ 是每一期财富结果的非线性凹效用函数（幂效用函数）。两种结果所导致的效用改变量分别为 $U(1200)-U(1000)$ 和 $U(1150)-U(1200)$。一种可能是，投资者只关心退休时的效用水平 $U(1150)$，也就是说，在这种情形中，中间情形的效用是无关紧要的，只有 $U(1150)$ 才是重要的；另外一种可能是，投资者有一个基准效用水平 U_0，他只关心相对于基准点的收益和损失，这是心理账户（mental accounting）的一种形式，即众所周知的狭窄框架（narrow framing），因为投资者只关心股票市场的结果，而不是关心总财富的结果（如收入变化或房地产价值变化导致总体财富发生变化）。终生效用在这里为 $U_0+U(+100)+U(-50)$，这也增加了“对 $U(-50)$ 赋予比 $U(+50)$ 更高权重”的可能性，从而获利后再遭受等可能性损失的效用为 $U(+50)+\lambda U(-50)$（其中 $\lambda>1$），这就是损失规避。损失规避对终生效用的影响很明显依赖于投资者监管投资组合的频率（如果你对你的投资组合频繁地进行估值（如每个月一次），与每 5 年

www.55188.com最好的股票论坛理想在线证券网欢迎您!

进行一次估值的情形相比，你会观察到更多的负的股票收益（Rabin 和 Thaler（2001）)）。

另外一种情形是，收益和损失的度量不是采用绝对值的形式，而是采用相对于初始效用的形式，随着财富的增加或减少，投资者缓慢地进行调整（这就是用相对于财富的“习惯水平”来对收益或损失进行度量的思想），这一思想在 Compbell - Cochrane 的消费习惯持续性模型中出现过（第 14 章），Babaeris、Shleifer 和 Wurgler（2001）也将这种思想应用到财富赌博之中（参见第 19 章）。事实上，上述构造在很多文献中是和幂效用函数一起使用的。

18.3　噪声交易者的存活

如果我们面对这样一个市场，精明投资者把价格设定为基础价值，噪声交易者根据经验规则行事。这就产生了一个问题，噪声交易者如何在这个市场中生存。如果在价格高于基础价值时噪声交易者持有股票，那么精明交易者必须把资产卖给噪声交易者，这使得价格下降。当股票价格朝基础价值方向下降时，噪声交易者会损失资金或破产，而精明交易者（如果他们愿意）会以较低的价格回购股票。另一方面，如果噪声交易者持有价格低于基础价值的资产，那么精明交易者会从他们手中购买股票，当价格朝基础价值方向运动时，精明交易者就会获利。因此，净效应是噪声交易者遭受损失，他们会在市场中消失，剩下的就只有精明交易者。当这种情况发生时，股票价格就等于基础价值。

当然，如果一大群噪声交易者陆陆续续进入市场（又陆陆续续破产），那么在某个重要时刻，价格有可能偏离基础价值。但是，也有人会认为，不可能出现的情形是，当前面的大量噪声交易破产后噪声交易者还进入市场。然而，企业家总是认为在别人失败时自己会成功。反过来想，一些噪声交易者在短期内会取得成功，这会鼓励其他噪声交易者进行模仿从而进入市场。模仿者忽视了如下事实：成功的噪声交易者事实上承担了更高的风险，成功仅仅是因为运气（这就是幸存偏差）。

我们能否解释为什么在存在精明交易者的市场中仍然有一些噪声交易者能够获得利润？答案与潜在的羊群行为有关。没有哪个精明交易者知道其他精明交易者打算在什么时刻通过交易促使股票价格向基础价值方向运动。因此，在噪声交易者存在的市场中，任何复杂的交易策略都不会没有风险，可能出现情形的是，噪声交易者使价格进一步偏离基础价值，这会导致精明交易者遭受损失，因此风险规避型精明交易者不能通过套利交易来完全消除噪声交易者的影响。如果足够多的噪声交易者有着共同的嗜好，那么噪声交易者风险就变成了具有普遍影响的（系统性）风险。这种风险不可能通过分散化投资来加以消除，均衡时它也必须获得报酬或风险溢价。因此，噪声交易与平均收益是一致的，这个平均收益比基础变量所给出的收益高。如果噪声交易者持有大量的股票，股票都包含噪声交易者风险，那么他们会获得比平均收益还要高的收益，从而在市场中存活下来。例如，如果噪声交易者积极地交易“不景气公司”的股票，那么这可以解释为什么这类公司的股票会获得比用 CAPMβ 进行风险调整后的平均收益还要高的收益。因为这类股票的噪声交易者风险高，所以相应收益也较高。

如果绝大多数投资者都遵从金融教科书的建议，消极地持有股票组合，那么噪声交易者对价格的影响会更大一些。如果噪声交易者根据“预感”加入到某一特殊群体，这个群体所持有的股票份额相同，那么市场组合的持有者不会有任何改变（除非噪声交易者的移动是如此的大以至于每种资产的“市场价值”比例需要发生变化）。与噪声交易者行动相反的是精明的套利者在市场中积极行动。在极端情形中，如果所有投资者都持有市场投资组合，但有一个噪声交易者进入市场的目的是希望购买某家特殊公司的股票，那么这家公司的股票价格会上升到无穷大。

套利交易者不仅要对基础变量进行预测，也要分散一部分精力去关注噪声交易者需求的变化。如果噪声交易者对某种股票很乐观，那么套利者通过投资银行的证券买卖活动来创造出

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

更多的证券（如高风险低价格债券、.com 公司、电信公司、共同基金、石油股票等）是值得的。假设某个企业集团对石油市场很感兴趣，噪声交易者被“石油”迷住了，那么企业集团将石油部门分离开来并发现新的石油股票是值得的，这些石油股票的价格异常的高，因此在噪声交易者中非常受欢迎。

套利者也可能像噪声交易者一样试图挑选噪声交易者所喜爱的股票，套利者的需求没有必要与噪声交易者的需求相反。正如企业家投资赌场来赚赌徒的钱一样，精明交易者花费大量的资源来收集噪声交易者未来需求可能变化方面的信息（如上世纪 90 年代的 .com 公司）是值得的。因此，一些套利者存在模仿噪声交易者行为的激励。例如，如果套利者发现噪声交易者是正反馈交易者（positive feedback trader)，那么当价格被推到高于基础价值时，套利交易者会顺应潮流，希望在价格最高时卖出股票。因此，他们有“放大热点”的作用。套利者希望长期价格能回复到基础价值的水平，但他们在短期内会“跟随趋势”，这一证据与如下事实一致：当套利者和噪声交易者跟随短期趋势时，短期（如周或月）收益具有正的自相关性；当一些套利者采取“长期观点”并卖出过高定价的股票时，长期（如三年或更长）收益具有负的相关性。另外，如果“消息”引发了噪声交易者的需求，那么这与价格对消息存在过度反应的事实一致。

群体行为

我们考察为什么市场中存在大量的噪声交易者，他们遵循简单的“经验规则”，容易受投资者情绪的影响（羊群效应)，他们不是根据基础价值的变化来采取相应的行动。为了使噪声交易者这个群体能对市场价格产生影响，他们的需求必须基本一致（如交易者的需求变化之间存在相关性）。

对“群体行为”的实验（Shiller（1989））包括 Sherif（1937）的“自动实验”(autokinetic experiment)，即要求个体在一个漆黑的环境里来预测光束的变动。实验中的个体对光束变动的幅度看法不一致（实际上光束没有变动）。当一群人参加同样的实验且每个个体都能听到其他个体的看法时，出现的结果是大家对光束变动的幅度有着一致的看法（但不同群体之间还是存在差异）。在 Asch（1952）的实验中，个体独立地比较线段的长短。先告知群体成员一个错误答案，然后让群体成员参与试验。个体单独参与试验时通常会给出正确的答案，但个体在面对群体压力时“频繁地”给出错误的答案。通过这个实验可以确定是，个体通常知道正确答案，但害怕与群体不一致。如果给定股票的正确价格或基础价值没有一个广泛接受的看法，那么投资者面临的是不确定性而不是风险（即他们不知道收益的真实条件分布），这就使得他们容易受交易者情绪所影响。

“观点扩散”（diffusion of opinion）模型通常不够准确。有证据表明，思想在长期内处于潜伏状态，然后由一些平凡的事件所引发。显然，新闻媒体在这里起着一定的作用，但对说服力的研究表明，家庭、朋友、同事之间面对面的非正式交流在信息传播过程中的重要性比新闻媒体大得多。这里有一套基于传染病模型的信息扩散数学理论。在这些模型中，“携带者”遇到一个“易感染者”就产生了一个“新的携带者”。携带者死亡的概率称为“死亡率”。如果传染病处于“传播状态”，那么传染病就会产生一个“驼峰形状”的模式。如果传染病不处于传播状态（因为传染率低，或易感染者的数量少，或死亡率高），那么新携带者的数量单调递减。将这个模型应用于投资者情绪的一个困难是，人们无法精确地量化模型中各个变量（如传染率）行为的决定因素，行为的这些确定因素在不同的情形中会有所不同。

Shiller（1989）运用上述思想所得出的结论是，20 世纪 50 年代和 60 年代的牛市与个体如何投资股票及个体之间股票收益（如投资俱乐部）差异的信息传播速度有关。他还注意到那段时期机构（比如养老基金）对股票的需求出现了增长，这种需求无法用个体为了使总储蓄保持为常数所出售的数量加以抵消。这是因为个体的股票持有数量不是均匀分布，那些购买养老金的个体没有股票可卖。

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

"羊群行为"或"跟随趋势"在房地产市场上可以经常看到，在 1987 年的股市下跌中（参见 Shiller（1990））以及外汇市场上（Frankel 和 Foot（1986）及 Allen 和 Taylor（1989））都可以观察到。Summers（1986）及 Shleifer 和 Summers（1990）表明，由模型（价格以某种方式持久地偏离基础价值）人工生成的股票价格时间序列确实能产生一个模拟实际价格行为（近似于随机游走）的时间序列，因此噪声交易者的某种持续行为与所观察到的数据基本一致。

当然，人们总是可以通过反复学习来调整自己的"经验规则"偏差。专家比普通人更（或不）容易产生偏差，减小这种偏差存在激励。然而，偏差似乎仍然存在，在结构发生变化的世界中，当学习是一个连续过程时，这种偏差肯定会持续下去。

18.4　异像

日历效应

周末效应是指股票价格从星期五收盘到星期一开盘之间会出现系统性下降的事实。对周末效应的一种解释是，公司和政府总是在星期一到星期五之间发布"利好消息"，而等到周末才发布"利空消息"。利空消息在星期一以"低"价格的形式反映出来。然而，在有效市场上，假设预期利润大于交易成本和风险偿付之和，一些投资者会意识到周末效应从而在星期五（以"高"价格）卖出股票，然后在星期一（以"低"价格）买进股票。这种行为可能会导致异像的消除。

所谓的一月效应是一种类似的现象。股票的每日收益在一月的前几天通常都是出奇的高。就美国而言，一种解释是，这是为了使资本损失投资者在年末卖出股票，这些资本损失可以抵消资本收益，从而降低税负。（在英国者有著名的"证券短期买卖"（bed and breakfasting）。）投资者在一月份购买股票，从而恢复原先所持有的投资组合。如果 EMH 成立，那么股票价格变化的这种可预测模式会导致免税交易者（如抚恤金）在 12 月价格低的时候买进股票，在一月价格高的时候卖出股票，这样就可以消除获利的套利机会。一月效应通常出现在一月的头 5 个交易日（Keane（1983）），并且似乎都集中在小公司股票上（Reiganum（1983））。在异像文献中还存在大量的其他日历效应（如年轮效应、节假日效应、月效应中的周效应）。问题是，分开考虑时在统计上显著的这些效应在同时进行考虑时是否仍然显著。

数据探测偏差

Sullivan、Timmermann 和 White（2001）注意到，大量的研究发现股票收益存在日历效应，并强调这些效应不是由"偶然"因素（数据挖掘）所导致的问题。他们使用 100 年的每日收益数据（S&P500）来考察 9500 多个可能的日历效应，这些日历效应包括文献中所没有报告的一些效应。（文献中所报告的日历效应都是"成功"的日历效应，这会使结果出现偏差。）基本思想是在所寻找 t 统计量的最大可能值基础上考察检验功效是否与其名义值不同。他们使用 *White*（2000）的"自助探测法"（bootstrap snooper），一旦考虑数据挖掘中所出现的效应，这种方法就能给出日历效应"成功"的 p 值。他们发现，最好的日历规则（要么从"大"要么从"小"的日历效应集合中进行选择）都无法保证（实际检验中）"可预测性"p 值的统计显著性，可预测性是针对平均收益或夏普比而言的。

此外，样本内的日历效应规则所导致的绩效在样本外时期中可能非常糟糕。因此，在对数据调查偏差进行考虑后，他们没有发现显著的日历效应。一般来说，数据调查思想试图考虑幸存偏好问题，因为研究人员没有发现的"预测性"结果没有公开发表（甚至没有投稿）。因此，我们在杂志上看到的论证可预测性的文章数目不成比例。

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

小公司效应

从 1960 年到 20 世纪 80 年代中期，在许多工业化国家中，小市值公司股票的平均收益比股票市场指数的平均收益高（Dimson、Marsh 和 Staunton（2002））。当然，根据 CAPM 模型，这可能是因为这些小公司股票的风险更高，从而 CAPM 模型的 β 值较大。然而，我们在第 8 章曾经注意到，小公司股票没有落在 CAPM 证券市场线上（Reinganum（1982、1983），Cochrane（2001））。从 20 世纪 80 年代后期开始，小公司效应这种异像似乎在学术领域或专业杂志上消失了，因此市场最终“消除了”这种异像（Cochrane（2001），Dimson、Marsh 和 Staunton（2002））。

赢家诅咒：“买入瘦狗，卖出明星股”

我们已经注意到，长期（3～5 年）股票收益存在均值反转现象。EMH 有效的一个关键问题是这种“可预测性”在扣除交易成本和对风险进行调整后能否保证超额利润的存在。DeBondt 和 Thaler（1985）的学术论文就是这种方法的一个例子。他们选取 1928 年 1 月～1932 年 12 月期间的 35 只“赢家”股票和 35 只“输家”股票（以 NYSE 普通股票的月度收益数据为基础）来构造两个截然不同的股票组合，然后观察它们在随后 5 年（“检验期”）中的绩效。重复上述程序 46 次，开始日期依次向前推进一年。最后，在 35 家公司权重相等的条件下，他们计算出了“检验期”的平均绩效（相对于 NYSE 指数的超额收益而言）。所得到的结果如下（图 1）：

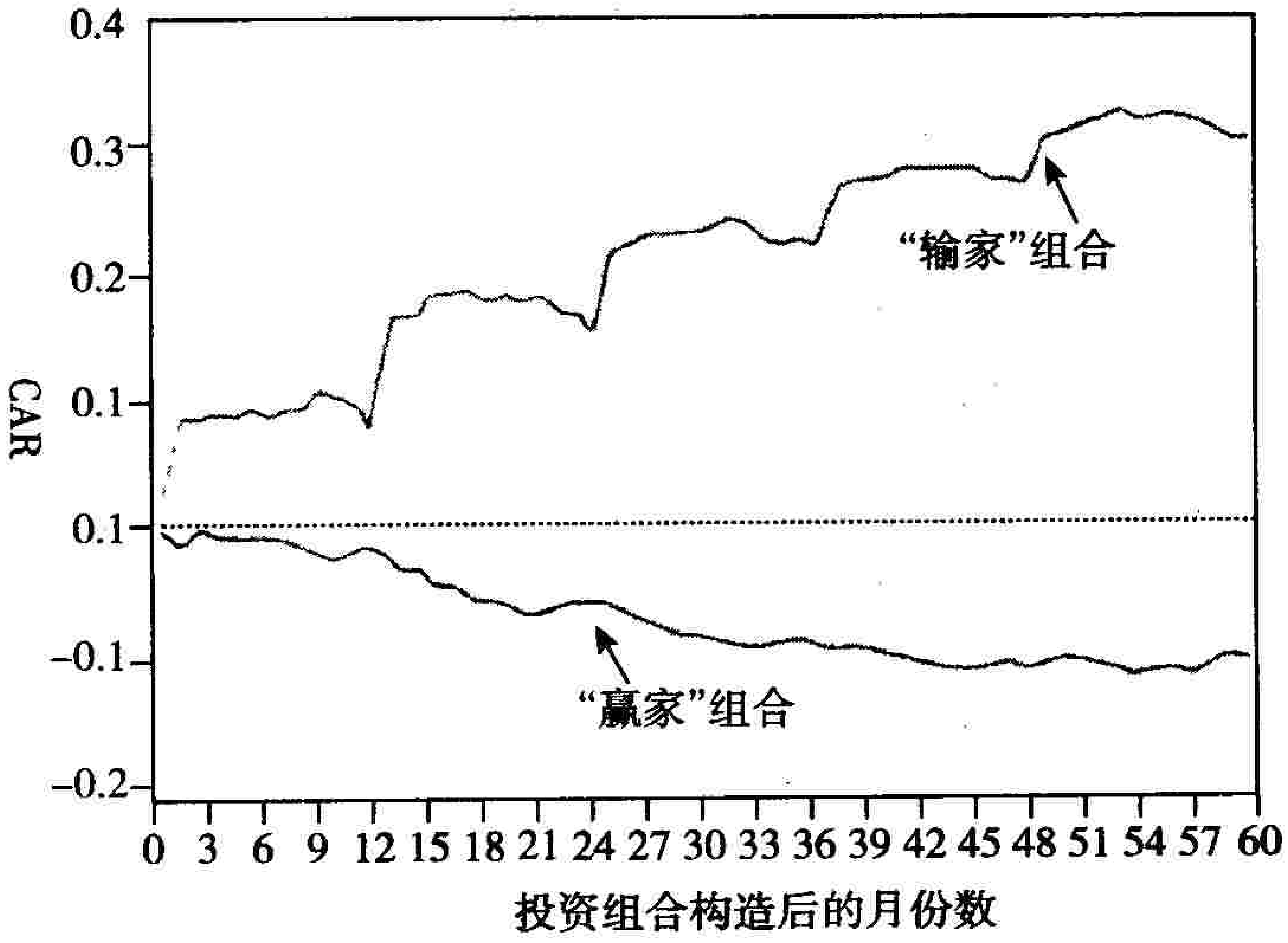


图 1　“赢家”组合和“输家”组合的累计超额收益

资料来源：DeBondt 和 Thaler（1989）。美国经济协会授权复印。

（ⅰ）“输家组合”的 5 年期收益为 30%，而“赢家组合”的损失约为 10%；

（ⅱ）“输家组合”的超额收益倾向于在一月发生（一月效应）；

（ⅲ）均值反转现象在长期中存在（即价格上升之后紧接着是价格下降，价格下降之后紧接着是价格上升）。

值得强调的是，所谓“输家组合”（即价格在过去出现过急剧下降）实际上是未来收益很高的股票组合，这与“输家”的定义在某种程度上是矛盾的。买入“输家组合”和“卖空赢家组合”的套利策略（即“买入瘦狗”和“卖出明星股”的策略）能击败持有 S&P500 的被动策略（参见 Debond 和 Thaler（1987））。

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

Bremer 和 Sweeney（1991）发现，上述结果在很短的时期内仍然成立。例如，对一天下降10%以上的“输家组合”而言，之后5天的收益为3.95%。他们只使用了大公司的股票数据。可以确信的是竞买竞卖价差占股票价格一个较大的百分比（这可能是对结果的一种扭曲）。另外，他们回避了“小公司效应”问题（即小公司具有“更高的”风险，所要求的均衡必要收益也更高一些）。因此，Bremer 和 Sweeney（1991）似乎也发现了超额利润存在的证据。

其他研究也证明了长期反向策略（contrarian strategies）能获得利润，这些研究包括：Chopra、Lakonishok 和 Ritter（1992）（对美国进行了研究）、Richards（1995、1997）（使用的是国内股票市场指数数据）、Lakonishok、Shleifer 和 Vishny（1994）。

上述结果的一种解释是，“输家组合”的“认知风险”较高，持有这些组合的必要收益也较高。早期研究对风险的度量较为粗糙，后来的研究经常用 Jensen 或夏普比来度量风险。

孪生股票

Froot 和 Dabora（1999）对 Royal Dutch 和 Shell 公司进行了研究，两家公司同意以60:40的比例合并现金流，但它们仍然是两家独立的公司。如果价格等于基础价值，那么 Royal Dutch 和 Shell 两家公司的股票价格之比为 RD/S = 1.5。但是，人们发现的价值高估或价值低估的程度分别为35%和15%，定价误差持续存在。这里几乎没有或有很小的基础变量风险，卖空交易不是很困难。当然，噪声交易者风险仍然存在，这（与短期理性交易者）是错误定价持续存在的原因。因此，不存在“免费午餐”，但“价格也不正确”。然而，价格在2001年却是“正确的”，两家公司的股票价格比为1.5，错误定价最终消失了。另外出现的情形是，在美国交易的 Royal Dutch 公司的股票和在英国交易的 Shell 公司的股票分别与 S&P500 指数和 FTSE 指数同步变动，这使得“基础”现金流与价格没有联系。这意味着两个市场都存在系统性噪声交易者风险，换一种稍微不同的叙述方式是，对特殊股票而言存在“偏好习性”现象。

美国存托凭证

美国存托凭证（ADR）是美国金融机构以信托方式持有的、在纽约进行交易的外国股票。这些股票也在“国内”市场上进行交易，但交易价格与在纽约的交易价格不同。因此，两者近乎是完全替代型证券，如果在纽约交易的 ADR 相对于国内市场存在一个溢价，那么套利者可以卖空 ADR 购买国内股票。但是，即使持有 ADR 的美国交易者在“国内”市场上也没有任何限制条件，两者之间仍然存在持久的价格偏差。这种异像应归因于噪声交易者风险和套利限制。

指数效应

当某只股票被编入（排除）S&P500 指数时，价格通常会有很大的上升（下降）。因此，价格变化与（能观察到的）基础价值变化不相关。换句话说，“没有消息的价格变动”出现了，这与 RVF 和 EMH 相违背。Wurgler 和 Zhuravkaya（2002）认为替代证券最为糟糕的股票（即套利时风险最高的股票）存在最大的价格跳跃。事实上，很难找到“好的”替代证券，因为 Zhuravkaya 发现，存在“最好”替代证券的股票收益的通常小于25%。因此，这种异像可能要归因于系统性基础风险的存在（和风险规避）。

封闭式基金

封闭式基金发行的基金单位份数是固定的，投资者之间可以对封闭式基金单位进行交易。基金筹集到的现金用于购买“一揽子”证券（通常是股票和债券的一个组合）。封闭式基金的基金单位可以在公开市场上进行交易。因此，基金的价值，即净资产价值（NAV），应等于基金单位的市场价格，但经常出现的情况却不是这样。

封闭式基金单位的价格通常与 NAV 不同。封闭式基金在刚设立时以高于 NAV 的价格进行

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

销售；在终结时以低于 NAV 的价格进行出售；在基金的存活期内通常以 NAV 的某个折扣比例进行交易。平均而言，折扣为 9 折，折扣比例随着时间会发生变化。另外，不同类型封闭式基金的折扣比例倾向于同步变动。这与 EMH 相违背，因为投资者可以以折扣价购买基金单位，同时卖空基金所持有的股票组合。因此，投资者最终能获得一个与折价比率相等的无风险利润。

封闭式基金会出现折价有几个方面的原因。第一，如果封闭式基金在基金升值后卖出证券，那么封闭式基金持有人就负纳税义务（资本所得税）。潜在的纳税义务使得基金价格低于标的证券的市场价值。第二，封闭式基金中的某些资产可交易程度很低（即“稀薄”市场交易）。第三，管理费用等代理成本也可以解释封闭式基金的折价现象。然而，Malkiel（1977）发现，封闭式基金的折价程度远比上述原因所解释的折价程度高得多，而 Lee、Shleifer 和 Thaler（1990）却发现，封闭式基金的折价程度主要由小公司股票的收益来决定。

这里还存在一种异像。在首次公开发行时，封闭式基金单位要支付承销成本，因此基金单位与真正的市场价值相比存在一个溢价。封闭式基金单位的价格通常在六个月内变动到折扣价格。令人奇怪的是，为什么投资者愿意在首次公开发行时购买基金单位（通过未来资本损失来支付承销成本）而不是等到六个月后以折扣价格购买基金单位呢？

行为金融学方面的思想与这类事实是一致的。如果噪声交易者风险是系统性风险，那么这种风险就获得能“定价”，因此封闭式基金的价格低于 NAV（参见下一章中 DeLong、Shleifer、Summers 和 Woldmann（1990b）中的模型）。在基金偿还时，所有噪声交易者风险都会消失，价格等于 NAV。在初始时刻，基金单位价值高估的原因是，在企业家创设基金时噪声交易喜欢这种类型的基金。不同类型基金折扣比率同步运动的原因是（假设基金的潜在现金流这个“基础变量”不会有很强的相关性）噪声交易者对所有类型基金的情绪通常会同步变动。Bodurtha、Kim 和 Lee（1993）发现，在美国交易的、持有德国股票的封闭式基金，与美国市场之间的联动程度比与德国市场之间的联动程度强。这再次证明基金现金流与价格之间存在分离，于是 EMH 遭到了违背。

Lee、Shleifer 和 Thaler（1991）也发现封闭式基金和小公司股票基本上被大量（美国）个体投资者持有，封闭式基金的折价比率与小股票收益之间有着很强的正相关性，这意味着噪声交易者的行为具有系统性。最近，Gemmill 和 Thomas（2002）构造了一个双变量 VECM 模型，模型中投资者的现金流与折价比率之间存在协整关系（因此两者在长期内存在相关性），非均衡误差修正项影响的是“现金流”（即 Granger 原因），而不是折价比率的变化。以 1991—1997 年期间 158 只英国股票型基金的月度数据为基础的分析结果为投资者情绪对折价的影响（对不同类型的基金而言）提供了更为深入的证据。

股票收益的联动效应

在有效市场中，收益的联动效应是由现金流（或贴现率）的联动效应所导致的，现金流的联动效应是由影响未来收益（例如，出售西班牙足球联赛的转播权会导致皇家马德里和巴塞罗那足球俱乐部的未来收益出现联动效应）或贴现率的共同因素所导致的。然而，正如我们在 Royal Dutch 和 Shell 公司股票收益情形中所看到的那样，根据 EMH，两种收益应该完全正相关，但事实并非如此。

为了反驳有效市场假说，Barberis 和 Shleifer（2003）注意到，“噪声交易者”对小市值股票、银行股等现金流几乎不相关的股票存在“偏好习性”（preferred habitat）。然而，如果噪声交易者的情绪具有系统性，那么这些“类别”收益之间会表现出很强的联动效应，类别中的新股与类别之间的联动性比以前强得多。Barberis、Shleifer 和 Wurgler（2001）认为，当某只股票被编入 S&P500 指数时出现的就是这种情形，因为这只股票相对于 S&P500 指数的 β 值会增加，但排除在 S&P500 指数之外的股票相对于指数的 β 值会减小。这些观察到的变化似乎与现金流的联动效应没有关系，因为“股票被编入 S&P500”这一事件没有传递现金流相关性方面

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

的任何信息。

最优分散化?

假设在投资者进行投资组合选择时均值方差理论是“正确”的方法（或模型）。人们发现，美国和日本的投资者选择国内资产的比例都很高（通常会超过 90%），选择国外资产的比例很低。事实上，包含人力资本的投资组合选择模型认为投资者应卖空国内股票，因为国内股票的收益与人力资本之间存在很强的相关性。这就是本土偏好（home bias）问题。

Grinblatt 和 Keloharju（2001）及 Huberman（2001）也发现了本土偏好问题，投资者总是偏好地理位置与他们最为接近公司的股票。Grinblatt 和 Keloharju（2001）发现芬兰投资者持有和交易的主要是芬兰的股票，而 Huberman（2001）却发现投资者持有本地“区域性贝尔营运公司（RBOC）”的股票数量比持有其他地区的“区域性贝尔营运公司”的股票数量多。

Benartzi（2001）发现 401 只养老金的股票持有计划具有非常强的本地偏好，即使“所拥有公司”的收益与个体劳动收入高度相关，情形也是如此。在安然（Enron）公司垮台后，人们发现安然公司的大部分职员购买的是那些主要持有安然公司股票的养老基金，不幸的是，他们也遭受了巨大损失。Benartzi 和 Thaler（2001）发现个体采取的分散化策略经常是简单的 $1/n$ 法则。个体把储蓄的 $1/n$ 分配到每一种投资选择上，不考虑股票和债券之间的不同配置效率。举一个例子来说明这个问题。个体可以在以下三种情形中进行选择：（A）“股票加债券型基金”；（B）“股票加平衡型基金”（平衡型基金持有 50% 的股票和 50% 的债券）；（C）“债券加平衡型基金”。在每一种选择中，个体对两种资产的持有比例为 50:50。运用 $1/n$ 法则来选择（A）就得到一个有效的股票 - 债券配置（S，B）=（50%，50%），选择（B）就得到（S，B）=（67%，33%），选择（C）就得到（S，B）=（33%，67%）。当参与者实际所给的选择就是（A）、（B）和（C）时，他们会采取近似公平的 50：50 法则，（A）、（B）和（C）中股票的平均配置比例分别为 54%、73% 和 35%。

根据 401 只基金（主要是股票型基金）中的 1/n 法则，个体最终会在股票上配置更大比例的资产。Benartzi 和 Thaler（2001）对 170 个退休储蓄计划进行了考察，根据股票型基金中股票配置比例将这些退休储蓄计划分为“低”、“中”、“高”三个类别。例如，如果使用的是 $1/n$ 法则，那么我们期望出现的结果是：在“低”类别中（基金所持有的股票比例相对较小（债券的持有比例较大一些））股票的配置比例很小。他们发现在“低”、“中”、“高”三个类别中股票的实际配置比例分别为 36%、65% 和 85%（在你的“计划”中，如果股票（相对于债券而言）的持有比例越高，那么你最终会持有更多的股票（即使从理论来看，股票的任何配置比例都是可能的））。

根据完全信息下的理性行为，我们很难解释上述事实。从心理学实验的结果来看，人们不喜欢这样的情形：他们不知道可能结果的客观概率分布。因此，他们对本土股票或自己工作的公司的股票更加“熟悉”。这是“模糊规避”的一个例子，你会规避你对结果的概率分布没有信心的情形。

买进和卖出

检验 EMH 的一种直接方法是考察“积极”投资者在重复交易中是否做出过“好的”投资决策。Odean（1999）发现投资者的平均投资收益比基准策略低。另外，对所有投资者而言，所买入股票的年平均收益比所卖出股票的年平均收益低，投资者更有可能卖出那些相对于初始购买价格而言价格上涨而不是下跌的股票。另外，频繁交易的投资者（通常是男性多于女性）获得的平均收益较低。简而言之，那些频繁交易的投资者都不是善于选股的投资者，他们最好是被动地持有分散化基准投资组合。也可能出现的情形是，主动型选股者和频繁交易者太过于自信，过于自信的新投资者群体陆续进入市场。他们不愿意卖出价格下跌的股票（通常称其为处置效应（disposition effect）），因为他们不愿意承认自己所犯的错误（根据收益和损

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

失的联合效用函数，这是理性的，参见 Barberis 和 Thaler（2003））。

在房地产市场同样可以观察到这种处置效应，人们在熊市中以低于初始购买价格的价格亏本卖出资产，他们设定的价格比具有类似特征的小商贩高 30%，小商贩是不会进行亏本出售的（因为他们会提前购买），损失规避型售卖者获得的价格实际上比商贩高（Genesove 和 Mayer（2001））。

投资风格

财经媒体和投资专家提倡多种投资风格。在过去十年中，最为流行的两种投资风格是“价值型投资”（value stock，也称为“价值增长型”策略）和动量策略（momentum strategy）（但它们是如何定义的呢）。两种策略之间的关键差异在于，构造投资组合的时期不同，持有投资组合的时间也不同。“价值增长型”投资策略买进“长期的输家”（“瘦狗”），卖空“长期的赢家”（“明星股”）（因此，这是一种反转策略），这种策略的绩效取决于长期均值反转效应（我们在第 4 章的统计检验中已经注意到这一点）。通过单变量自回归或 Fama - French 长期收益方程或方差比统计量就可以得到均值反转的证据。另外，均值反转也可能是因为长期收益可预测性的存在，预测变量（如红利/价格比、账面市值比）本身也具有均值反转特征。

价值增长型策略

DeBondt 和 Thaler（1985）提倡买进“瘦狗”的策略，“瘦狗”指（比方说）由过去四年中价格下降幅度最大的 1/10 股票所构成的投资组合，投资组合的持有时间至少为一年。这是“价值增长策略”的一个例子，这种策略与单变量统计检验结果一致。Jegadeesh 和 Titman（1993）及 Fama 和 French（1996）也给出了类似的例子，但对价值策略的定义却不同。例如，“瘦狗”的一种定义是高账面市值比（BMV）股票。当然，如果“瘦狗”的 CAPM β 值很大，那么这对 EMH 有利。不幸的是，这种情形事实上并不会出现。正如我们已经看到的那样，Fama 和 French（1996）使用平均（横截面）收益发现，“瘦狗”组合（和其他的 1/10 股票组合）的平均收益可以用 HML（high - minus - low）β 来进行解释（也就是说，Fama - French 的“瘦狗”有着高的 HML β 值）。Fama 和 French（1996）发现做多“瘦狗”卖空“明星股”所获得的平均收益为 0.75% ~ 1.6%。但是，“买入瘦狗”仍然可以产生比市场投资组合更大的夏普比，尽管每单位风险的收益每年都不一样。

Lakonishok、Shleifer 和 Vishny（1994）及 Ali、Hwang 和 Trombiey（2003）发现，在每年的七月卖空 *BMV* 最低的 1/5 美国股票（根据账面市值比进行分类时要对规模进行调整），买进 BMV 最高的 1/5 股票（每年都要进行调整），这种策略在后续一年、两年和三年中的平均收益分别为 8%、21% 和 30%。为了使用这种策略，你必须有充足的资金和一定的信用额度来承担所遭受的损失，也得有卖空能力。这种策略仍然是存在风险的策略。基于账面市值比的解释是，市场参与者低估 BMV 高的股票未来收益和高估 *BMV* 低股票的未来收益的行为导致了错误定价（La Porta、Lakonishok、Shleifer 和 Vishny（1997），Skinner 和 Sloan（2002））。但是，为什么套利者不利用这种机会来进行套利从而消除错误定价呢？

风险与错误定价

因此，BMV 效应所导致的长期收益可预测性是因为策略的事前风险太高还是因为存在长期错误定价？这是一个很难回答的问题，因为这需要一个可接受的风险模型，也需要对交易成本进行估计（因为这是一种长期投资策略，所以交易成本不会太高）。Ali、Hwang 和 Trombiey（2003）试图对这一问题进行回答。他们对美国 1976—1997 年期间 BMV 1/5 组合三年期平均收益的横截面特征进行了考察，用多种方法来度量交易成本和投资组合的风险。（每年都重复这种横截面回归，使用 Fama 和 MacBeth（1973）程序来检验解释变量的统计显著性，同时对非白噪声误差项进行 New - West 修正）。

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

交易成本可以用平均交易量、（平均而言的）竞买竞卖价差百分比［＝（竞买价－竞卖价）/0.5（竞买价＋竞卖价）］和日收益为0的频率（即非流动性股票）来进行度量。卖空成本（如果经纪人要求直接得到股票收益，那么卖空挤压（short squeeze）的可能性就存在）用机构投资者持有股票的比例来表示。机构投资者持有的股票比例越小，股票借出方遇到卖空挤压的可能性就越小，套利成本就越高。投资者的其他复杂度量指标还包括股票的分析师数目和股票的机构投资者数目。

度量套利风险的简单变量是残差的方差，残差是用过去250天的每日收益数据来对市场模型进行估计时所得到的。这里的思想是，特殊风险相对较高的股票不大可能得到正确定价，因为套利者的投资组合大多不是良好分散化投资组合，他们的投资组合中只有很少几只股票。因此，这种特殊风险对套利者来说很重要，因为我们无法对这种风险进行套期保值（这与系统风险不同），投资期限有限的套利者不大可能持有特殊风险很高的股票，因此这种定价“异像”无法消除（Shleifer和Vishny（1997））。“方差变量”在统计上显著（即使横截面回归模型中还包含其他变量），这意味着套利功能受到了限制。（换一种稍微不同的叙述方式）Ali、Hwang和Trombiey（2003）发现，收益波动性最高的10%股票给出的三年收益达到51.3%（投资组合包含*BMV*最高的1/5股票组合的多头和BMV最低的1/5股票组合的空头）。波动性最小的1/10股票组合的相应收益却只有1.7%，这种模式在22年时期中的20个年份里出现。当然，套利风险的这种解释不是决定性的，因为如果高BMV变量确实反映了金融危机风险的增加，那么收益高度波动的那些公司遭受这种风险的可能性就会更大。

对“价值”的绝大多数定义而言，基于“价值策略”的任何投资组合的构造都存在很大的重叠。例如，对低价格、低市盈率、高BMV和低市场价值（“规模”）等而言，每一个1/10组合里的公司都存在很大的重叠（对最高的十分之一组合和最低的1/10组合而言，重叠尤其严重），因为驱动所有这些变量的因素才能获得定价。换句话说，即使账面市值比在大量研究中都有很好的表现，使用价值的一种定义而不使用另外一种定义所带来的解释能力增加或可预测性增加都非常小。当然，在APT框架下，你会接受“HML β 是度量风险的有效变量”这一结论。因此，“买进‘瘦狗股’能获得比市场更高的收益这一事实”根本就没有任何问题，因为这些股票有很高的“BMV风险因子”β 值。换句话说，Fama－French三因子模型（即市场 β、SMB β 和 HML β）可以解释基于账面市值比进行分类的横截面收益。事实上，我们在前面就注意到HML β 占横截面解释力的很大一个部分。

可以肯定的是，在任何收益期内，试图“击败市场”的任何策略都是高风险策略。对“价值策略”而言，你必须买进价格在几年内一直下跌的股票，这种类型的许多公司大都面临“危机”甚至濒临破产。与此同时，你必须卖空几年内价格连续上涨的股票，这种类型的公司是当前优质公司的典范。此外，你可能在“不景气的时候”买进“瘦狗”，又不得不持有1到4年，在经济复苏的时候，你有可能要遭受损失。这种策略适合于勇敢的投资个体，如果个体被投资银行雇佣，在策略带来正的收益之前，他可能已经被解雇了（例如，在2000年网络股泡沫破裂后，那些在1998年和1999年就主张卖出股票的分析师还在为自己进行辩护，但他们也失去了工作）。在这种情况下，也许真实的风险比事后度量的风险高；因此，价值投资策略的超额收益实际上仅仅是承担额外风险的回报，更不用说交易成本和卖空方面的高昂成本。

牛市行情

为了强调受限套利的可能性，先问问自己，在观察到美国和英国自1995年以来的牛市后你在1999年都做了什么。假设你通过基本面分析得出S&P500目前被高估的结论（Shiller（1999））。你应该卖空S&P500指数，买进一个类似的“正确定价的”指数（如小公司股票的“Russell 2000”指数），然后进行等待。如果S&P500中大公司的任何“消息”与小公司的任何消息之间不存在很强的相关性，那么你有可能面临基础风险。你还会面临噪声交易者风险（S&P500甚至会继续上升）和模型风险（例如，相对于基础变量而言，你真的确定S&P500被

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

高估了吗?)。如果你持有该头寸6个月或1年，那么它的价格可能会急剧下降。当然，如果你能一直持有到2002—2003年，那么你的策略可能被证明是正确的，但到那个时候，你的精神可能已经崩溃或者你已经失业了，或者两种情形都发生了。此外，你可能听取了John Ralfe的大胆策略，他是Boots（英国药店和化学家）养老基金的基金经理，他在2001年将所有资金从股票市场转移到债券市场，这从事后来看可以说是过去50年来最好的择时交易。(事实上，投入长期债券包括指数联动（index - linked）债券可能是为了追求养老计划的时间与基金现金流之间的匹配)。

动量策略

投资组合的构造时期在这里都很短，通常基于上一年（或更短时期）的收益来构造投资组合，持有期也很短，通常只有一年。动量策略是买进短期的“赢家”，卖出（卖空）短期的“输家”。你买进过去几年上涨幅度最大的“1/10”股票的组合，通过卖出过去几年中下跌幅度最大的“1/10”股票的组合来为购买进行融资（也称其为相对强势操作策略（relative strength strategy））。这种策略每月重复一次，每个投资组合的持有期为一年。Fama和French（1996）报告的结果是，在美国使用动量策略所获得的平均月度收益为：1963年7月到1993年12月期间为1.3%（每年为15.6%）；1931年1月到1963年2月为0.38%（每年为4.56%）。显然，动量策略成功击败市场的时期只有1963—1993年，这说明动量策略不一定可取。

最近，Jegadeesh和Titman（2001）在1993年文章的基础上对动量策略重新进行了考察。他们使用NYSE和AMEX股票但将NASDAQ股票排除在外（也将股价低于5美元的股票排除在外）。在每月的月末，根据过去6个月的收益将股票分为10组，然后用每组中的股票来构造等权重投资组合，每个投资组合的持有期为6个月。我们用一个例子来简单说明“基于前6个月的排序”。12月份的“赢家”组合由第6到11月份收益最高的10%股票组成，而11月份的“赢家”组合由第5到10月份收益最高的10%股票组成。以此类推，8月份的“赢家”组合由第2到7月份收益的最高收益的10%股票组成。这些投资组合是“重叠”的，最长的“回顾时间”实际上为一年。动量投资组合是一种自融资策略，因为投资组合要求卖空“输家”组合$P10$（由过去六个月中收益最低的10%股票组成）买进“赢家”组合$P1$（由过去六个月中收益最高的10%股票组成）。这些投资组合每月进行一次调整。1965—1997年期间“赢家”组合$P1$的平均收益为每月1.67%，而“输家”组合$P10$的平均收益为每月0.58%。因此，“赢家”组合比“输家”组合的收益每月高1.09%（统计量为5.65），结果在1965—1989年和1990—1997年期间也成立。

为了评价动量策略（相对于预期收益的）实际收益绩效，他们将$P1$到$P10$这10个投资组合中每一个的收益分别对超额市场收益R_m或Fama - French“三因子”收益R_m、R_{SMB}和R_{HML}进行时间序列回归来得到Jensen α值。对市场收益因子而言，$P1$的α值最大（0.54%*p.m*），$P10$的α值最小（-0.57%*p.m*），两者之间每月1.1%的差异在统计上是显著的（$t=4.59$）。类似地，对Fama - French三因子模型而言，α值之间的差异为1.29%，这在统计上也是显著的（$t=7.8$）。因此，$P1 \sim P10$动量策略似乎能产生统计显著的异常收益。

Jegadeesh和Titman（2001）对“持有期之后时期上”的动量投资组合收益（即第13到第60个月上的收益）进行了考察。$P1$到$P10$组合的累积动量收益在一年后达到最大值12%（年度收益，平均每月为1%）。在此之后的收益几乎为负，60个月后的累计收益仅为1%（每年）。这就是长期均值反转现象。

上述结果与Barberis、Shleifer和Vishny（1998）的“保守主义偏差”模型一致。在他们的模型中，个体在对先念信息进行更新时新信息的权重要低一些，价格对新信息的调整非常缓慢。一旦新信息被完全吸收后，股票收益就不再具有可预测性，因此持有期之后（即12个月以后）的超额收益为0。这个结果与“自我归因偏差”（self - attribution bias）模型一致

www.55188.com最好的股票论坛理想在线证券网欢迎您!

(Daniel、Hirshleifer 和 Subrahmanyam (1998))。在后面一个模型中，精明投资者将事后短期“赢家”归因于优秀的投资技能，将事后“输家”归因于坏的运气。这种过于自信意味着他们会购买更多的“赢家”，慢慢地将股价推到高于基础价值的地方。理性交易者最终会进入，“赢家”组合的收益最终也会出现反转。Hong 和 Stein (1999) 在解释短期动量和长期收益反转的模型中假设存在两组投资者。噪声交易者根据历史收益来进行决策，而理性交易者使用“基本面消息”来进行决策。但是，信息的传递存在滞后，所以价格在开始时对消息反应不足，这就导致了动量收益。噪声交易者此时观察到了动量收益，然后将过去“赢家”组合的价格推到高于基础价值的地方，这最终使价格出现反转。每一类投资者都理性地使用信息，但流向噪声交易者的信息存在滞后，不同类型的交易者使用不同的信息集合（即“红利方面的消息”和历史收益)。我们将在下一章中详细地讨论这些模型。

尽管数据与这些行为模型一致，Jegadeesh 和 Titman 还是注意到，持有期之后负的股票收益主要发生在投资组合构造后的第四年，四年的滞后期似乎太长了一些。另外，绝大多数收益反转发生在一月份，而 6 个月持有期的动量利润都发生在一月之外的月份中。这种现象特别不对称。

总的来说，Jegadeesh 和 Titman (1993) 所描述的成功动量策略在 1990—1997 年的新数据中仍然成立，这使得“1993 年的早期结果归因于数据探测偏差”的证据有所减弱。有待进一步考察的是这些利润在 2000—2002 年的股票市场中是否依然存在。

Daniel 和 Titman (2000) 对上述方法进行了发展，他们用三种分类方法对美国股票收益进行分类 (1963 年 7 月 ~ 1987 年 12 月)。NYSE、AMEX 和 NASDAQ 的所有股票根据市值（“规模”)、账面市值比（“价值”）和先前一年的收益（“动量”）可以分为 5 类，由此产生了 125 个 (5 × 5 × 5) 投资组合。这显然对 Fama 和 French (1993) 的两种分类方法（“规模”和“价值”）进行了扩展。

然后，我们使用如下策略：买进（高价值，高动量）组合，卖出（低价值，低动量）组合。我们称其为“HH - LL”策略。这种“分类”方法能产生一个非常漂亮的结果，即 34 年中有 31 年的收益为正，平均月收益为 1.04%（每年 12.64%)，这个收益显著地不为 0 ($t = 5.66$)。HH - LL 组合的 CAPMβ 为 -0.258，Jensenα 为每月 1.17% ($t = 6.6$)。于是，HH - LL 组合具有明显的异常收益，能提供针对市场收益风险的保险。

另外，在将月度收益 R_{HH-LL} 对共同基金（超额）收益指数 R_{MF} 或 $R_{MF} - R_M$ 进行回归时，(Jensen) α 都为正数，约等于 1.1（在统计上显著)。前者意味着，如果把 HH - LL 组合添加到基金之中，那么夏普比会所有增加。后者意味着，添加 HH - LL 组合能增加预期收益，但不会增加共同基金的跟踪误差（相对于市场投资组合而言)。夏普比的增加取决于 α_i^2/σ_i 的大小，这里的 α_i 是截距，σ_i 是“新”的超额收益 (new excess return) R_i 对所选择的（超额）市场收益 (chosen excess market return) 进行回归时残差的标准差 (Jobson 和 Korkie (1984))。他们还证明了如下结论，如果 HH - LL 组合包含 2% 的共同基金组合，那么“扩大后的”共同基金组合的平均夏普比 $\{SR_{aug} = (\bar{R}_p - \bar{R}_f) / \sigma (R_p - R_f)\}$ 将从 4% 上升到 12%（即上升了三倍）——尽管这可能要对 HH - LL 组合中某些证券进行卖空。跟踪误差的夏普比 $\{SR_{trk} = (\bar{R}_p - \bar{R}_m) / \sigma (R_p - R_m)\}$ 由 0.12 上升到 1，这意味着，当在共同基金中加入 HH - LL 方向上“2% 的误差组合”时，风险收益绩效会更好一些。考察 2000—2002 年熊市时期 HH - LL 策略是否继续有效是很有趣的一件事情。

动量策略的基础是在短期收益中所发现的微弱正自相关性，月度收益的自相关系数为 $R_{sq} = 0.0025$ (1/4%)。Cochrane (2001) 注意到，对个体股票收益而言，$\sigma = 40\% p.a$，因此在 10 个投资组合中最高平均收益为

$$E < R \mid R > x = \frac{\int_x^{\infty} r f(r) \, \mathrm{d}r}{\int_x^{\infty} f(r) \, \mathrm{d}r}$$

由正态分布可知$\int_x^\infty f(r)\,\mathrm{d}r = 0.1$意味着$x = 1.2816\sigma$。因此，10个投资组合中最高的平均收益$E<R \mid R>X>$比均值大1.76个标准差。假如均值收益为10% p.a，$\sigma = 40\%$ p.a，则收益最高组合与最低组合的平均收益分别为80%和-60% p.a。然后，他指出上述研究中连续使用动量策略（但时间很短）的月度只有1%。为了看清这一点，注意个体股票收益的月度标准差为$\sigma_{mth} = 40\%/\sqrt{12} = 11.55\%$，预测收益的月度标准差为$\sigma(\hat{R}) = \sqrt{R_{sq}}\sigma_{mth} = \sqrt{0.0025}(11.55\%) = 0.58\%$。因此，在正态分布假设下，最前面的1/10组合将获得比所有股票平均收益高$1.76 \times 0.58\% = 1\%$的月度收益。“买入赢家，卖空输家”的动量策略应获得2%的月度收益，而它在最好的时候也只能获得1.5%的收益（见上面）。换句话说，最前面的1/10动量组合的价格在过去一年中会上涨80%左右，而输家组合则会下降60%左右。因此，只要收益之间存在很弱的正相关性，结果就会导致下一年1%的平均月度收益，但这个1%不是异常收益（Cochrane（2001））。

交易成本

人们在考察动量股票时发现，收益主要来自于11~12月中卖空流动性差的股票。交易成本有可能很高，有可能抵消掉前面所注意到的超过收益（Carhart（1997），Moskowitz和Grinblatt（1999））。在Jegadeesh和Titman（1993）中，动量交易策略的利润用平均交易成本来进行调整，因为平均佣金率和市场影响效应存在（Berkowitz、Logue和Noser（1988））。但是，平均交易成本没有考虑交易成本的横截面差异（小公司股票有较高的交易成本）和时变性。平均成本也不包括竞买竞卖价差、税收和卖空成本。动量策略属于密集型交易，每6个月要交易4次（买进-卖出，卖出-买进），因此6个月中的平均异常收益为6%（即α的估计值），如果每笔交易的成本超过1.5%，那么这个利润就会被完全抵消。

Lesmond、Schill和Chunsheng（2004）对上述交易成本进行了详细考察（考虑了不是所有动量组合都每6个月进行一次“交易”的事实），结果发现，对所有交易成本进行调整后，美国股票的动量策略没有获得利润。他们使用30%~70%断点来研究Jegadeesh和Titman（1993，2001）及Hong、Lim和Stein（2000）中的绩效结果。P_1（最优绩效组合）和P_3（最差绩效组合）的交易成本分别为4.5%和5%（每6个月），而Jegadeesh和Titman（1993，2001）所使用的平均值为0.5%。这些“新的”交易成本意味着，$P_1 - P_3$策略每6个月的平均动量收益（6%）很难弥补交易成本（扣除交易成本后的收益在统计上通常不显著）。有待进一步研究的问题是，对其他文献中所记录的成功动量策略（如美国以外的市场、债券市场和其他部门中的动量策略）的交易成本进行详细考察后动量收益是否为零。

实证文献中的异像肯定对EMH提出了质疑（由于股票收益存在可预测性，人们可以获得异常利润）。然而，必须注意的是，在上述异像中，很多异像主要发生在小公司和“处于困境中的”公司（如一月效应和规模效应、价值增长策略和动量策略、封闭式基金折价）。可能出现的情形是，事前认知风险比事后度量的风险高，因此，这些异像在对认知风险进行调整后可能不会产生超额收益。有些异像会一直持续下去，有些异像会消失（如小公司效应）。异像文献肯定使EMH陷入了艰难境地。如果你勤于观察并愿意在重复博弈中承担风险（在通往金融极乐世界的道路上损失一些钱财），那么你就有可能在特定时期中捕捉到获利机会。即使在对风险（和交易成本）进行修正后，你仍然可以获得超额收益，但在我们周围仍然不存在令人目眩且显而易见的“免费午餐”。通过选股来击败市场是很难做到的，尽管我的MBA学生在20世纪90年代的牛市中不相信这一点，但在2000年以后，他们倾向于接受这一观点。

18.5　公司金融

管理层在什么时候发行新股呢？在 EMH 下，发行股票的“时机选择”没有任何优势可言。市场价格总是等于公司现有项目和计划投资项目的利润现值（PV）。但是，如果对未来投资的现值持乐观态度的系统性噪声交易者存在，那么这会在高的股票价格中有所反映，管理层应很好地利用价格被非理性高估的事实来选择合适的新股发行时间。然而，管理层不应将资金投入到固定投资项目上去，因为新固定投资项目的“真实”NPV 值为 0（在均衡状态下）。这些资金可以投资到 NPV 为正的其他公司上，或者直接留存为现金。相反，管理层应回购价格被非理性低估的股票，但固定投资仍然没有发生变化。上述论点在现实中可能需要进行调整。然而，如果新募集的资金没有用到新投资项目上，那么噪声交易者就会迫使管理者出局。另外，经理人可能会在扩展权力的基础上进行大量投资，而不是为了股东的财富最大化进行投资。

Baker 和 Wurgler（2000）发现，高账面市值比公司不会用新募集的资金来增加投资，募集资金的目的是为了增加现金。当股票的市场价格很高时，新股发行的比例也较高（相对于总股本和新债券发行的总和而言），这个反映了市场时机选择的观点。总之，从公司横截面来看，账面市值比是公司新股发行的一个很好的预测指标。高账面市值比公司会发行更多的新股，而低账面市值比公司则回购股票。

如果噪声交易者对两家完全相同的公司（如规模、行业、有形资产的比例等）具有不同的情感，两家公司的账面市值比不同，那么择时交易的思想意味着两家公司有着不同的权益/负债比率（即不同的资本结构）。与此一致的是，Baker 和 Wurgler（2002）发现，从公司的横截面来看，公司历史上最高的账面市值比是“权益资本在总资本中的比例”的良好预测指标。调查证据也支持股本发行时机选择的观点，因为超过 65% 的 CFO 都说他们是在市场高估股票时决定发行股本的。

尽管上述证据与噪声交易者思想一致，资本结构的逐序规则还是与理性假设基本一致。假设管理者认为，“高的”股票价格应归因于理性投资者的交易，因为投资机会的 NPV 为正。如果内部资金（和债券融资）已经耗尽，那么管理者的理性行为是在价格“高”时发行更多的股本。这与噪声交易者情形相同。然而，两者之间的关键差异在于：在理性情形中，管理者肯定会进行更多的固定投资，而在噪声交易者模型中（管理者理性）则不会有其他投资（NPV <0），管理者参与噪声交易“游戏”的情形当然要排除在外。因此，两者在预测方面没有显著差异，两个故事之间的差异很有可能存在很大的问题，也存在很大的争论。

非理性经理人

让我们放松经理人理性的假设，替代的假设是他们对自身能力过于自信。Roll（1986）的并购活动“傲慢假说”认为，并购不会有内在的协同收益，它之所以发生是因为并购发起人过于自信，过高地估计了协同收益。于是，被并购企业的价格上升，而并购发起人的价格则下降相同的数量。因此，有很多并购实际上没有产生协同效应（有证据证明这是正确的（Demodoran（2000）））。

经理人的乐观主义也可以解释金融投资项目中的逐序观点。如果乐观的经理人认为股票市场持久地低估了他公司股票的价值，那么他会借助于债券市场和或使用内部资金来为公司扩张进行融资。因此，固定投资与现金流的关系可能强于托宾 Q 理论中的关系（有证据证明这个结论是正确的（Cuhbertson 和 Gasparro（1993）））。

18.6 小结

- 行为金融与“不存在免费午餐”的观点一致，但认为这并不意味着“价格是正确的”。
- 如果套利者属于风险规避型且基础风险具有系统性，那么在近似替代证券（对股票而言）不存在时，持久的错误定价就会发生。另外，在不存在基础风险时，只要套利者属于风险规避型且期限较短，噪声交易者风险具有系统性，那么错误定价就会发生。
- 心理学研究的结论为行为金融提供了证据。
- 异像文献发现人们在很多情形中可以获得异常利润量。在很多情形中，击败市场的策略（如周末效应、一月效应、价值增长型策略和动量策略）要求投资者持有风险头寸，这些利润在扣除交易成本和对风险进行调整后还是否存在，人们对此还存在争论。
- 其余的异像（如 ADR、孪生股票、指数效应、封闭式基金）似乎一直存在，几乎不涉及风险，因此这些实证证据为行为金融观点提供了强有力的支持。

第19章

行为模型

目的

* 说明理性交易者和噪声交易者之间的交互作用如何导致与内在价值不一样的均衡价格以及理性交易者如何导致过度波动性。

* 考察短期主义如何导致错误定价。

* 说明投资者在收益方面可获得的公开或私人信息的错觉如何导致收益公告后股票价格的持续反应（post - earning announcement drift）、动量效应和股票价格的过度反应。

* 分析噪声交易者（如动量交易者、风格投资者）与基本面交易者之间的交互作用如何导致股票价格的不足反应（之后就是过度反应）和不同类别股票之间的交叉相关性。

* 说明非标准效用函数（在损失规避型投资者的跨期优化模型中，这种效用函数是消费和“其他变量”的函数）如何解释股票收益中的“典型事实”（如过度波动性、可预测性和股权溢价之谜）。

至此，我们已经在相当一般的情形中讨论噪声交易者的影响，现在是讨论更为正式模型的时候了。一类模型假设噪声交易者和理性交易者之间的交互作用导致了使市场出清的均衡价格。另一类模型假设所有代理人都是同质的，但偏好不是标准偏好（效用取决于财富和消费水平的变化，与收益相比，代理人更害怕遭受损失）。这些模型经常在“标准方法中”假设投资者追求跨期效用最大化。

19.1　简单模型

Shiller（1989）对噪声交易者所能解释的经典实证结果进行了简单分析。理性套利者（smart money，即市场上“聪明的钱”）对股票的需求比例为 Q，Q 大致可以根据均值方差模型来确定，

$$Q_t = (E_t R_{t+1} - \rho) / \theta \tag{1}$$

如果 $E_t R_{t+1} = \rho$，那么理性套利者的股票需求为 0。如果 $Q_t = 1$，那么理性套利者持有股票，这时的必要预期收益为 $E_t R_{t+1} = \rho + \theta$。因此，$\theta$ 是吸引理性套利者持有所有股票的风险溢价。

我们现在假设（Y_t / P_t）等于噪声交易者所持有的股票份额。在均衡状态下，理性套利者和噪声交易者所持有的股票份额和应等于 1，

$$Q_t + (Y_t / P_t) = 1 \tag{2}$$

将（1）代入（2），

$$E_tR_{t+1}=\theta\left[1-(Y_t/P_t)\right]+\rho \tag{3}$$

因此，理性套利者的预期收益依赖于他们所认为的噪声交易者当前和未来需求，噪声交易者的需求越高，当前价格就越高，理性套利者对收益的预期就越低。由（3）和如下定义

$$E_tR_{t+1}=E_t\left[(P_{t+1}+D_{t+1})/P_t-1\right] \tag{4}$$

就可以得到

$$P_t=\delta E_t(P_{t+1}+D_{t+1}+\theta Y_t) \tag{5}$$

其中，$\delta=1/(1+\rho+\theta)$。因此，重复前向替代就得到

$$P_t=\sum_{t=0}^{\infty}\delta^t(E_tD_{t+1}+\theta E_tY_{t+1}) \tag{6}$$

如果理性套利者是理性的，清楚地知道噪声交易者的需求，那么理性套利者就可以计算出市场出清价格，它等于内在价值（E_tD_{t+1}）和噪声交易者未来需求 E_tY_{t+1}的加权平均。这个模型的缺点是，噪声交易者的需求完全是外生的。然而，正如我们在下面即将看到的那样，我们仍然可以从这个模型中得到一些有用的直觉。

如果 E_tY_{t+1}和噪声交易者的总需求是随机的且均值为 0，那么（6）中 E_tY_{t+1}的移动平均值对 P_t 几乎没有影响，P_t 主要受基本面所主导。价格会偏离内在价值，但这种偏离是随机的。另一方面，如果噪声交易者需求具有持久性（例如，在大的 Y_t 值之后是更大的值），那么噪声交易者需求的微小变化对目前的价格都会产生很大的影响，这可能使得价格在很长的一个时期内极大地偏离内在价值。

Shiller（1989）使用上述模型来说明，在 EMH 不成立时，基于收益对 t 时刻已知的信息变量进行回归的市场效率检验，拒绝 EMH 的功效为什么这样的低。假设红利（和贴现率）为常数，因此 EMH（没有噪声交易者）意味着股票价格为常数。现在假设市场实际上完全由噪声交易者所驱动，假设噪声交易者的需求为

$$Y_t=\mu_{t-1}+\mu_{t-2}+\mu_{t-3}+\cdots+\mu_{t-n} \tag{7}$$

μ_t 是白噪声。方程（7）有如下性质：t 时刻的单位需求增加使得 Y 在未来时刻的变动有一个平方峰（square hump），平方峰在 n 期后消失。使用（6）就可以得知，价格变动（$P_{t+1}-P_t$）仅仅是因为人们对噪声交易者未来的需求预期进行了修正，修正的方法是用 δ，δ^2，δ^3 等来进行加权。因为 $0<\delta<1$，所以价格变化主要受 μ_t（而不是历史值 μ_{t-j}）控制。然而，因为 μ_t 是随机的，所以模型中的价格变化（根据构造，它由噪声交易者所主导）基本上是不可预测的。

Shiller 使用（6）针对 Y_t 中持续性的不同取值（由滞后长度 n 来给出）及 ρ 和 θ 的各种取值来生成 ΔP_{t+1}，然后将所生成的 ΔP_{t+1}对只包含 P_t 的信息集合进行回归。在 EMH 成立的条件下，我们预期回归中的 $R^2=0$。当 $\rho=0$，$\theta=0.2$，$n=20$ 时，Shiller 发现 $R^2=0.015$。小的 R 值支持收益为常数时的 EMH，但这一结果可以由如下模型来导出：模型中的价格变化完全由噪声交易者决定。另外，即使价格变化很难预测，价格水平相对于内在价值而言仍然有很大的偏离。他所计算的结果还有，如果所生成的数据包括 4% 的常数红利价格比，那么收益 R_{t+1}对价格红利比 $(D/P)_t$ 回归的“理论 R^2”等于 0.079。因此，“收益与 t 时刻的信息只是弱相关”的经验证据与“价格由噪声交易者（而不是基本面）来决定”的结论不一致。总之，Shiller 对经验证据给出了中肯的意见。使用实际数据所发现的证据不是股票收益不可预测（如 EMH 所建议的那样），而是股票收益可预测的程度很低。然而，后一个结论也与噪声交易者起着重要作用的模型结论不一致。

如果 Y_t 的行为是外生的（即与红利无关），但又是平稳和均值反转的，那么我们预期收益是可预测的。在比 Y 的平均值高的那些地方，后面必然会出现 Y 的下降（向长期平均值方向运动）。因此，价格是均值反转的，根据历史收益可以对当前收益进行预测。

另外，这个简单的噪声交易者模型可以解释红利价格比和下期股票收益之间正的相关性。如果红利随着时间的演变只有很小的变化，那么由 E_tY_{t+1}的增加所导致的价格上升会使红利价格比下降。如果 Y_t 是均值反转的，那么价格在将来会下降，因此收益 R_{t+1}也会下降。因此，

正如实证研究所发现的那样，$(D/P)_t$ 与收益 R_{t+1} 是正相关的。Shiller 也注意到，如果噪声交易者的需求 Y_{t+1} 受历史收益（即惯性效应，bandwagon effect）或历史红利的影响，那么与（6）中第一项（价格反应中的内在价值部分）固定的情形相比，股票价格可能对当前红利产生过度反应。

19.2　噪声交易者行为的优化模型

在 DeLong、Shleifer、Summers 和 Waldmann（1990b）的“漂亮”模型中，理性套利者和噪声交易者都属于风险规避型，他们的目标都是追求最终财富效用的最大化。期限为某个有限值，套利存在风险。在（基本）模型中，只有噪声交易者风险，没有基本面风险（即红利是已知的）。噪声交易者对风险资产的狂热追求为他们自己和理性套利者带来了风险。理性套利者在价格变化正确分布的基础上形成对未来价格的最优预测，但噪声交易者的预测存在一定的偏差。噪声交易者对价格的错误认知程度 ρ_t 等于噪声交易者的预测值与最优预测值之差，

$$\rho_t \sim N\ (\rho^*,\ \sigma^2) \tag{8}$$

如果 $\rho^* = 0$，那么（平均而言）噪声交易者的预测值与理性套利者的预测值是一致的。如果噪声交易者的普遍态度是悲观的（即处于熊市之中），那么 $\rho^* < 0$，股票价格低于内在价值，反之亦反之。即使人们认为 ρ^* 在长期内会偏离最优预测值，“消息”也会产生，乐观主义者和悲观主义者都存在短暂的异常变化，变化大小为 $\rho_t - \rho^*$。ρ_t 的设定可以很随意，但基于内省（introspection）和行为/群体实验证据有直观吸引力。

在 DeLong 等的模型中，股票的内在价值为常数，我们可以很随意地将期设定为 1。市场中只有两种资产：风险资产和无风险资产。噪声交易者和理性套利者都是风险规避型，都具有均值方差偏好，因此他们对风险资产的需求与预期收益正相关，与噪声交易者风险（参见附录 I）负相关。噪声交易者的需求也取决于他们所认为的股票价格（即变量 ρ_t）是处于牛市行情中还是处于熊市行情中。风险资产的供给是固定的（等于 1），市场出清时的均衡价格为 P_t。确定 P_t 的方程看起来很复杂，但我们可以将其分解成几个部分，从而得到一些直观感觉：

$$P_t = 1 + \frac{\mu}{r}\rho^* + \frac{\mu}{1+r}\ [\rho_t - \rho^*]\ \ - \frac{2\gamma\mu^2\sigma^2}{r\ (1+r)^2} \tag{9}$$

其中，μ 是噪声交易者占总投资者的比例，γ 为无风险（实际）利率，r 为（绝对）风险规避系数，σ^2 是噪声交易者对价格错误认知的方差。如果不存在噪声交易者，那么 $\mu = 0$，（9）意味着市场价格等于基础价值（为 1）。现在假设我们处于某个特殊的时点上，噪声交易者对股票价格的长期预期与理性套利者一致（$\rho^* = 0$），“意想不到的情形”不会发生（不存在异常的牛市和熊市），因此 $(\rho_t - \rho^*) = 0$。

现在，我们认为噪声交易者对未来价格的预期与理性套利者相同。然而，均衡市场价格仍然不是内在价值的反映，市场价格小于内在价值（方程（9）中右边最后一项）。这是因为噪声交易者风险存在，他们的潜在行为会对价格产生影响。因为市场价格低于内在价值，作为对噪声交易者风险的补偿，理性套利者（和噪声交易者）能获得一个正的预期收益（即资本利得）。这种错误定价可能是这个模型最为重要的结果，模型涉及价格与内在价值之间的持久偏离。我们称（9）中第三项为“基本的错误定价”效应。

现在来看（9）中的第二项，我们知道，如果噪声交易者认为长期股市是牛市（即 $\rho^* > 0$），那么他们会将价格推高到大于内在价值的位置。第三项反映的是短期价格的异常上升或下降。这两项说明，在一个特殊的时期里，价格可能高于或低于内在价值。

如果 $\rho_t - \rho^*$ 是均值为 0 的随机变量，那么实际价格会随机地偏离“基本的错误定价”水平。在这种情形中，股票价格可能存在过度波动性（内在价值的相对波动性为 0）。根据（9），价格的方差为

$$E_t\ (P_t - E_tP_t)^2 = \frac{\mu^2 \mathrm{Var}_t\ (\rho_t - \rho^*)}{(1+r)^2} = \frac{\mu^2\sigma^2}{(1+r)^2} \tag{10}$$

因此，过度波动性越强，噪声交易者对价格错误认知的方差 σ^2 就越大，市场中噪声交易者的比例 μ 就越大，借贷资金的成本 r 也就越低。为了使模型能复制价格运动的持续性和观察到的股票价格大幅上升或下降的现象，我们需要引入“偏好”和“潮流”等概念。从宽泛意义上讲，这意味着牛市之后就是熊市，我们可以用随机游走 ρ_t^* 来进行表示

$$\rho_t^* = \rho_{t-1}^* + \omega_t \tag{11}$$

$\omega_t \sim N\ (0,\ \sigma_\omega^2)$。（注意，$\sigma_\omega^2$ 与前面方程中的 σ^2 不同。）投资者在任意时点的最优预测值 ρ^* 都等于当前值。然而，当“消息” ω_t 到达时，噪声交易者会改变对 ρ_t^* 的看法，这种“认知变化”会一直持续下去。必须明白的是，（9）中第二项的随机游走 ρ_t^* 意味着 P_t 在长期中会来回变动，从而 P_t 具有“牛熊轮回”模式。

为了说明问题，Fortune（1991）假设 ω_t 和（$\rho_t - \rho^*$）都服从 *niid*，在（9）中使用 r，μ，γ 的代表性值来生成时间序列 P_t（如图 1 所示）。图 1 表明，就拟合结果而言，价格下降到内在价值 85% 的位置，价格在短期内会出现急剧上升和下降的现象。

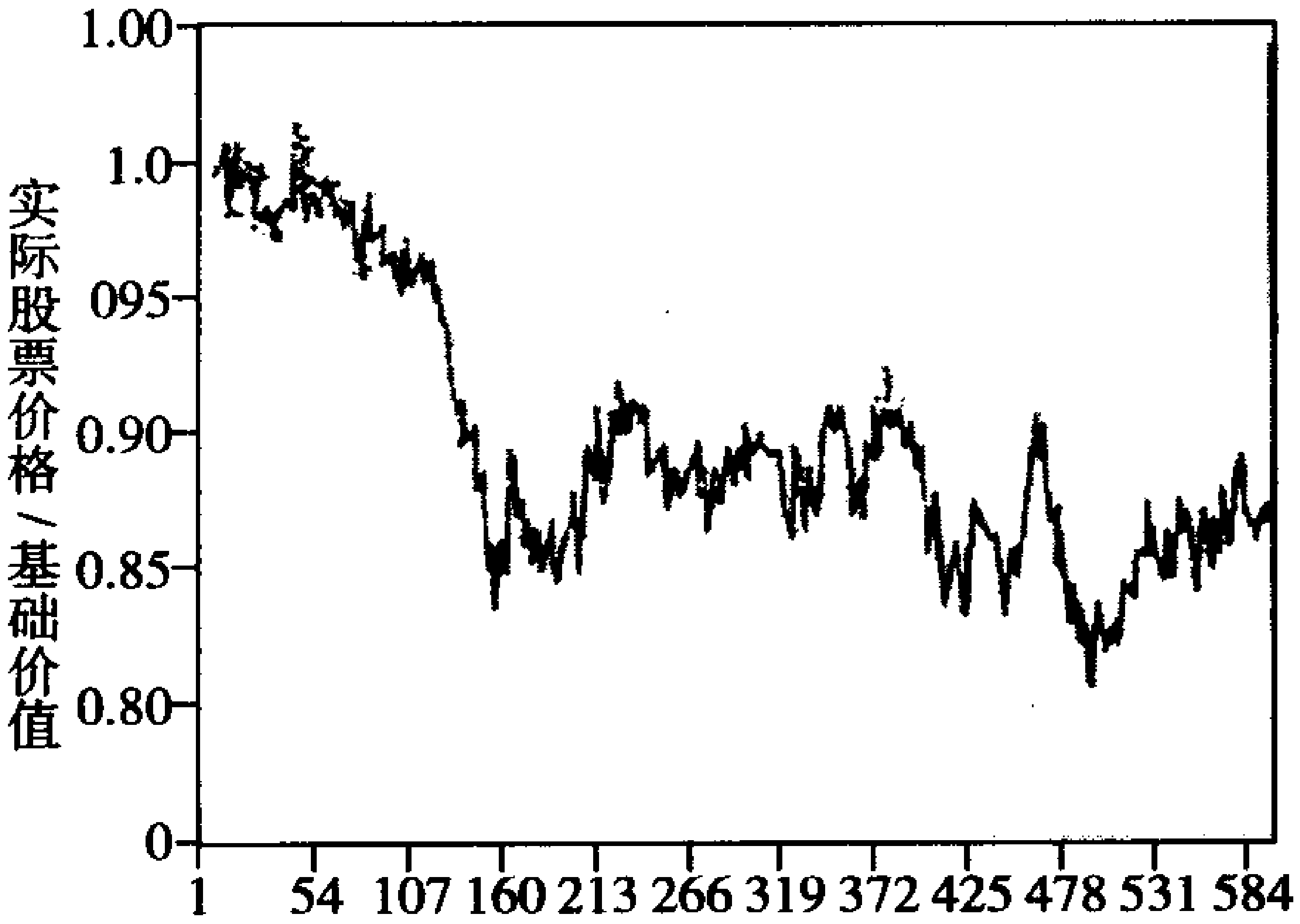


图 1　所拟合的 50 年股票价格

资料来源：Fortune（1991），图 6，*p.* 34。波士顿联邦储备银行授权复印。

假设 σ^2 具有自回归性质（如 ARCH 和 GARCH 过程），就可以在模型中为价格引入其他持久性来源。假设“从理性套利者到噪声交易者的‘转换率’（conversion rate）具有周期性”也是不合理的，这使得 μ（噪声交易者的比例）具有持续性，从而 P_t 也具有持续性。在这个模型中，因为套利不完备（incomplete），所以价格在大部分时间中与内在价值不相等。（$\rho_t - \rho^*$）的持续性也具有均值反转性。这意味着，价格是均值反转的，根据历史收益或红利价格比等变量，可以部分地预测股票收益。

噪声交易者能否存活

DeLong 等证明了如下结论：噪声交易者的比例在每一期中都是固定的（即 μ 为常数），即使噪声交易者倾向于买高卖低，（尽管无法保证）他们存活下来依然是完全可能的（反之亦反之）。这是因为他们过于乐观，低估投资组合的真实风险。因此，他们在行情看好时倾向于“持有更多的”风险资产。另外，如果噪声交易者风险 σ^2 很大，那么理性套利者不会花费太多精力去购买价格低估的资产。

如果假设理性套利者（s）转变成噪声交易者（n）的比率取决于噪声交易者所获得的（大于理性套利者的收益）超额收益（$R^n - R^s$），那么我们就可以在模型中引入模拟（imitation）

概念。

$$\mu_{t+1} = \mu_t + \psi \left(R^n - R^s \right)_t \qquad (12)$$

其中，μ 位于 0 和 1 之间。DeLong 等也在模型中引入了基本面风险。每期的风险资产收益是随机变量 $r+\varepsilon_t$，$\varepsilon_t \sim N\ (0,\ \sigma_\varepsilon^2)$。在这个版本的模型中，噪声交易者存活的概率始终大于 0。这是因为他们所谓的“创造空间效应（create space effect）”存在，风险增加得如此的高以至于风险规避型理性套利者不会为了任何潜在的收益而进行套利（这个结果要求 ψ 很“小”，否则新的噪声交易者会影响价格，这种影响可以根据已退出市场的“老”噪声交易者来进行预测）。

封闭式基金

我们已经注意到，封闭式基金倾向于折价销售，不同时间的折价比率不同，不同基金之间的折价比率也不同。这些基金有时也会进行溢价交易。使用我们的噪声交易者模型就可以解释这种“异像”。假设风险资产就是封闭式基金本身，而安全资产就是实际标的股票。理性套利者会在封闭式基金和标的股票之间进行套利交易（例如，如果基金进行折价销售，那么他们会买入基金，卖出股票）。然而，即使 $\rho_t = \rho^* = 0$，基金（风险资产）也会因为内在的噪声交易者风险（参见（9））而进行折价销售。噪声交易者的情绪变化（即 ρ^* 和 $\rho_t - \rho^*$ 中的变化）会导致不同时间的折价比率不同，噪声交易者风险也会演变成系统风险，大部分基金的折价比率同向变动。

在噪声交易者模型中，大量的封闭式基金倾向于在相同时刻启动，即在噪声交易者对封闭式基金情绪高涨的时候（即 $\rho^* > 0$，$\rho_t > 0$）进行销售。当已有的封闭式基金存在溢价时，理性套利者就会购买更多的股票（价格相当低），将其打包为基金投资组合然后再卖给那些乐观的噪声交易者。

债券价格的变化

从实证结果来看，如果债券的长短期收益差 $R-r$ 为正，那么长期利率倾向于下降，长期债券的价格会上升。这与利率期限结构的纯预期假说矛盾。“异像”的这种典型事实与我们的噪声交易者模型一致，长期债券在这里就是风险资产（短期债券为无风险资产）。当 $R_t > r_t$ 时，噪声交易者认为长期债券的价格异常的低。如果噪声交易者的偏好是均值反转的，那么他们预期债券价格在将来会上涨，从而长期利率下降，这就是我们在实证研究中所观察到的现象。当然，即使噪声交易者模型能解释这些典型事实，为了对债券市场的噪声交易者模型进行规范检验，我们还有很漫长的路要走。

总之，DeLong 等模型的主要特征是：均衡价格存在被低估的可能性。上面提到的其他效应依赖于噪声交易者的情绪变化，而情绪变化具有一定的持续性。尽管“持续性”从直觉来看具有一定的吸引力，但它仍然不是优化模型的结果。

短期主义

在只有理性套利者的市场中，一些投资者对收益采取短期主义观点，这不会导致价格偏离内在价值。观点的基础是向前递推的理性定价公式。如果你在今天（任意的时刻）买入而在明天卖出，那么你的收益（部分地）取决于预期资本利得和你明天所能获得的价格。但后者依赖于你的售卖对象在 $t+1$ 时刻所设想的 $t+2$ 时刻的股票价格，以此类推。因此，“以基本面为基础的理性”短期主义投资者所进行的计算与无限期界投资者相同。

如果投资者期限是有限的且噪声交易者存在，那么上述结论不成立。实际上，理性套利者的期限越长，他就越愿意基于价格与内在价值的偏差来进行风险套利，这是因为，他每期可以确定地获得红利支付，他有足够长的时间来等待价格上涨到高于内在价值的那一刻。然而，即使投资者的期限“很长”（但有限），他仍然面临转售价格风险。在“很长”的持有期中，分

红在总收益中的比例很高，但由于“最终价格”存在不确定性，投资者仍然会面临很高的风险。

从噪声交易者模型可以看出，如果噪声交易者的情绪对公司股票的影响越小（即降低 σ^2），那么价格被低估的程度就越轻，价格就会出现上涨。通过如下方式可以降低不确定性：

（a）增加当前红利（而不是将利润投资到不确定的长期投资项目上，如 *R&D* 开支）；

（b）债务与股本的互换；

（c）股票回购。

Jensen（1986）和其他学者的实证研究表明，（a）~（c）确实能使公司股票价格上升，这与我们所描述的噪声交易者影响的解释一致。因此，当噪声交易者存在时，人们预期资本结构的变化会影响公司价值（这与 Modigliani - Miller 假说矛盾）。

动摇理性交易者

在前面提到的 DeLong、Shleifer、Summers 和 Waldmann（1990*b*）模型中，理性交易者总是使价格向内在价值方向运动，但无法完全消除由“动量”噪声交易者所导致的定价偏差。假设“动量”交易者（即价格上涨时买入，价格下跌时卖出）存在，假设的文献证据非常之多。我们知道，那些图表主义者喜欢追逐“趋势”，在价格下跌后进行止损操作。投资组合保险也倾向于在价格下降（上升）后卖出（买入）股票。例如，如果你已经卖出一份看涨期权，那么你可以买入股票来为你的头寸进行套期保值（即 δ 套期保值）。如果股票价格连续下跌，那么期权的值也会下降，新的套期保值头寸要求你卖出一些股票。（严格地说，投资组合保险适用于用“股票 + 远期”组合来复制“股票 + 看跌期权”组合的情形，但我们的 δ 套期保值例子给出了类似结果。）因此，从逻辑来看，投资组合保险意味着你在股票价格下跌后卖出股票（在价格上涨后买入股票），参见 Cuthbertson 和 Nitzsche（2001b）。也有证据表明（Andreassen 和 Kraus（1990）），经济学专业的学生在面对真实的股票价格数据时倾向于在价格小幅上涨后卖出股票（反之亦然）、在价格连续上涨后买进股票（即动量策略）。

DeLong、Shleifer、Summers 和 Waldmann（1990a）使用上述分析框架，但允许理性交易者（他们意识到了噪声交易者的动量买入策略）采用动量交易策略。因此，理性交易者在股票价格上涨之后买入股票，希望驾驭这个波段，在价格重新回到内在价值之前卖出股票。因此，套利者的短期行为会受到动摇，甚至会导致短期收益出现更大的自相关性（在新的基本面信息到达之后），但长期收益是均值反转的（即负自相关性）。

由于篇幅限制，我们不可能在这里完整地叙述这个模型，我们只概述主要特征。模型中存在四个时期（0，1，2 和 3）和三种类型的交易者。理性交易者追求期末财富（消费）的最大化，拥有均值 - 方差资产需求（mean - variance asset demands，即与下期的预期收益成比例，与收益的方差负相关）。动量交易者的需求取决于前期价格变化，而“被动型”投资者的风险资产需求取决于最后一期的价格（即机械地低买高卖策略）。市场在每个时期内都会出清，基本面（即红利）方面存在一些噪声信号，这些信号会影响理性交易者的需求，从而影响他们对动量交易者未来行为的看法。在 $t=1$ 时刻，（第 3 期期末）红利方面的利好消息会导致价格上升，因为以基本面为基础的交易者存在。这会反过来导致 $t=2$ 时刻动量交易者需求的增加，在 $t=3$ 时刻，价格最终会回到内在价值。即使理性交易者并不认为动量交易者的未来需求会增加，过度反应仍然存在，但是，如果理性交易者也随大流，那么这种过度反应就会加剧。因此，进入该模型中的富有远见卓识的理性交易者数量的增加会导致“过度反应”加剧。

模型结果与“市场玩家”的文献证据一致。投资银行拥有客户买卖的“内部”信息，他们据此对动量交易者的需求进行预测（参见第 29 章市场微观结构）。这种形式的更强的一个变体是“提前行动”（front running），即造市商在执行客户买卖指令之前就用自己的账户买卖股票。（这引起了纽约 Attorney General 注意，Eliot Spitzer（2003））。在美国，联合投资组织（investment pools）经常会推荐一些热门股，以此来吸引动量投资者。Soros（1987）在 20 世纪

60年代到70年代的投资策略也可以视为是对未来的群体行为进行赌博，他在60年代的公司合并类型股票和70年代的房地产信托投资类型股票上进行了长期（成功的）赌博。1992年9月，他还成功地实施了卖出英镑的短期策略，他这次因为在几个星期之内赚取10亿美元而闻名于世（尽管这种策略是属于“追逐趋势”还是“基于基本面的偏离”还处于争论之中）。

FX预测服务公司的调查数据（Frankel和Froot（1988））表明，在20世纪80年代中期，美元升值（利率不变）使得预测者认为，一个月内美元会升值但美元在年末会贬值。尽管他们认为美元相对于内在价值而言已经被高估，但他们依然建议投资者在今天买进美元。DeLong、Shleifer、Summers和Waldmann（1990a）的模型表明，这种策略是一种完全理性的策略。

动量交易者是不是真的很愚蠢？他们会遭受损失吗？他们会被逐出市场吗？批评家们通过如下方式来回答这些问题：假设其他动量交易者（例如，使用新的技巧，比如神经网络、遗传算法等）进入市场，或已有动量交易者带着新的赌资回到市场。（如果你经常阅读《金融时报》或《华尔街日报》，你会发现这是经常发生的事情，尽管华尔街当前处于艰难时期，这在将来发生的频率可能要低一些。）另外，如果大量动量交易者在同一时期遭受损失，那么他们会宣称“每个人都做得很糟糕”，他们会坚持他们的投资理念。有证据表明，养老基金不会因为绝对损失而改变投资理念，但会因为比竞争对手遭受了更多损失而更换基金经理（因此，一些动量交易者仍然会在市场中存活下来）。另外，这种变化不仅由收益来决定，还受其他因素（如管理成本、分析师的研究贡献、投资风格等）的影响。DeLong、Shleifer、Summers和Waldmann（1990*b*）的早期模型也表明，噪声交易者可能比理性交易者承担了更多的市场风险，所以即使他们的判断是错误的，他们也能获得正的收益。

19.3　Shleifer - Vishny模型：短期主义

公司股票的价格被低估不是DeLong等噪声交易者模型的直接结果，因为这个模型要求噪声交易者的行为对所有股票而言具有系统性。然而，在Shleifer和Vishney（1990）的模型中，他们考察了高借入成本对个体股票错误定价程度的影响。他们发现，当前的这种错误定价在某些股票上尤其严重，这类股票的错误定价发生在更远的时期而不是下一个时期。假设实际投资项目的长期偿付是不确定的，项目通过发现股票的方式来进行融资，而股票的真实价值在一段时间之后才会在市场中显示出来。在Shleifer - Vishny模型中，这些股票的价格会被严重低估。于是，公司可能不愿意进行长期投资，即使投资项目可以获得利润，公司管理层的行为是一种短期主义。也就是说，他们宁愿选择获利少的短期投资项目，也不愿意进行长期投资，因为短期投资策略使得股票价格被低估的程度较轻，他们所面临的因竞争对手或董事会重组而失去工作的风险要小得多。这是对实物资源的一种错误配置。我们通过考虑无限期界情形来开始模型描述，模型中的理性套利者对真实价格在什么时候回到内在价值并不关心。

完美市场中套利利润的时机选择

如果理性套利者能无限制地进行借贷，那么她不在乎错误定价何时才能回到内在价值上来。在表1中，我们考虑价值低估的一种简单情形，r为借入成本，证券的基本回报（即分红回报q）为10%。如果被错误定价的证券在一期之后价格由5美元上升到内在价值6美元，那么一个时期中包含红利支付的价格为6（$1+q$）=6.60美元。在1期期末，套利者必须偿还本金和利息，即5（$1+r$）=5.50美元。如果价值在两期后才达到内在价值，那么套利者在$t+2$期能获得6（$1+q)^2=7.26$美元，但必须支付$t+1$到$t+2$期之间的利息。然而，就现值而言，无论错误定价在什么时候被根除，套利者始终能获得1美元的收益。另外，在完美资本市场中，套利者会使用任何套利机会，因为她能在任何时刻借入更多的资金。

表 1　　套利收益：完美资本市场

假设：

内在价值 = 6 美元

当前价格 = 5 美元

利率，$r=10\%$ / 每期

风险资产的收益，$q=10\%$ / 每期（相对于内在价值而言）

理性套利者以 10% 的利率借入 5 美元，在 $t=0$ 时刻购买股票

售卖价格（包括红利）	**偿还贷款**	**净收益**	**收益现值**（$r=10\%$）
时期 1　$6(1+q)=6.60$	$5(1+r)=5.50$	1.10	1
时期 2　$6(1+q)^2=7.26$	$5(1+r)^2=6.0-5$	1.21	1

在有限期限情形中，基本面风险和噪声交易者风险都可能使套利遭受损失。如果资金的提供者（如银行）发现他们很难对套利者选择价值低估股票的能力进行评估，那么他们就会对套利者实行借贷限额。另外，他们会对套利者收取更高的利息，因为他们对套利者的了解信息比套利者本身少得多（即信息不对称情形下的利息比信息对称情形下的利息高）。

在上述例子中，$r=12\%$，股票的基本面回报仍然为 10%。如果错误定价是立即而不是过一段时间才被消除，那么套利者就会获得更多的收益（参见附录 II）。如果对套利者施加严格的信用额度，那么他们就会承担额外的成本，也就是说，如果将资金与套利者的头寸相联系，那么他就无法使用潜在的套利机会。

套利者获得利润越多，所借入的资金就越多，持有价值低估股票的头寸就越大。因此，他会试图使（发信号）资金供给者相信，他非常“精通”短期套利交易，因为长期头寸很高且风险巨大。因此，理性套利者有动力去进行短期投资，而不是去消除长期套利机会。

Shleifer 和 Vishny（1990）的模型包含噪声交易者和理性套利者（参见附录 II）。短期资产和长期资产在未来某个时刻有着相同的支付，但短期资产的真实价值在市场中得以体现的时间比长期资产早。在均衡状态下，套利者的理性行为导致了长期资产在当前时刻更大的定价偏差，错误定价在较长时期之后才会表现出来。“长期”和“短期”这两个术语指的是定价偏差得以表现的时间（而不是这两种资产实际的偿付时间）。两种资产都存在定价偏差，但长期资产的定价偏差更大一些。

从本质来看，相对于错误定价证券的基本面收益而言，模型更多地依赖于套利者的资金成本。因此，套利者在清偿头寸（即卖出价值低估的证券）之前的等待时间越长，他所付出的成本就越高。他所获得的收益越高、偿还“昂贵”债务的时间越早，他的状况就越好。因此，在模型中重要的是“所承担的成本”或每期的资金借入成本。交易者对错误定价的长期资产的需求比错误定价的短期资产低，因此，前者的当前价格低于后者的当前价格。

未来偿付（利润）不确定的投资项目（在遥远的某个时刻增值）可以通过那些内在价值在很长一段时间都无法体现的资产来进行融资（例如，从英国到法国的海峡隧道工程在进行融资的许多年后，旅客收益才开始增加）。在这个模型中，这些资产严重被低估（相对而言）。

Shleifer 和 Vishny（1990）的另外一个结论与短期主义的观点不同，这个结论关注的是公司经理的行为。他们推测公司经理对股票定价偏差的权衡程度不同。相对于价值高估而言，他们认为同等程度的价值低估要更为糟糕一些。这是因为，价值低估会导致公司董事会更换经理，也可能导致恶意收购。另一方面，价值高估给经理所带来的福利很少，因为他们通常不会持有大量的股票，他们的收益与股票价格之间没有太强的联系。因此，经理没有激励去进行长期项目投资。

恶意收购者可能放弃长期投资项目，增加短期现金流或当前红利，所有这些都可能降低不

确定性，减少错误定价的持续时间。他可能将兼并过来的公司以更高的价格进行出售，因为当他取消长期投资项目后，价值被低估的程度会有所减小。上述情形意味着，一些可以获利的长期投资项目被牺牲了，因为投资者是短期主义者，他面临的是“高昂的”借贷成本，还有可能受到借贷约束。这与如下观点相反，恶意收购将那些无效率的经理（不是价值最大化的追求者）更换为效率更高的经理。因此，如果理性套利者没有足够的时间等待长期套利机会的到来，那么他们会支持恶意收购，因为恶意收购能降低错误定价的程度，能让他们迅速地了解套利交易头寸。

19.4　传染

Kirman（1993a）的“漂亮”模型与DeLong等的模型有很大的差异，因为前者直接考虑了个体之间的交互作用、个体观点受新成员影响而改变的速度、“羊群效应”和“传染”现象。昆虫学家注意到了“羊群”现象。当人们把蚂蚁放到两堆相同食物（食物要不断地得到补充）中间时，他们发现，两堆食物上分布的蚂蚁数目不对称。过一段时间后，80%的蚂蚁在吃一堆食物，而另外的20%的蚂蚁在吃另外一堆食物。有时出现的情形则相反。然后，他们重复进行试验，此时只放置一堆食物，同时安放两个可以到达食物的对称桥梁。在开始的时候，仍然是80%的蚂蚁经过一座桥梁，20%的蚂蚁经过另外一座桥梁，从直觉来看，人们所设想的情形是，蚂蚁以50∶50的比例通过每一座桥梁。蚁群吸纳新成员过程（recruitment process）的一种类型是所谓的“一前一后加入（tandem recruiting）”，在蚂蚁找到食物返回蚁穴之后，通过接触或释放某种化学物质来吸纳新成员。Kirman注意到，Backer（1991）记载了类似的羊群行为，当人们遇到非常相似的两家餐厅（价格、食物、服务等都一样，但位于同一道路的两边）的时候，绝大多数人选择在一家餐厅就餐，即使他们需要“排队”，情形也是这样。这里要注意的是，群体中的一部分可能存在外部性，我们假设蚁群没有这种外部性。

我们已经注意到，股票价格可能长期偏离内在价值。这种解释“‘吸纳成员’并导致一段时期内出现聚集然后‘离开’现象”的模型对观察投机性资产上的行为非常重要。Kirman指出，尽管经济学家（不同于昆虫学家）倾向于在优化行为的基础上进行选择，但优化不是存活的必要条件（例如，植物能存活下来，因为植物拥有进化系统，它们的叶子能吸收太阳能，如果植物有树干，那么它们会活得更好，因为有了树干，它们就能接触到更多的阳光）。Kirman的吸纳新成员模型存在以下几个方面的假设：

（ⅰ）对世界的看法只有“黑”和“白”两种情形，每个代理人在任何时刻只能选择一种看法。

（ⅱ）总共有N个代理人，系统由认为世界是“黑”的代理人数量（$=k$）决定。

（ⅲ）系统的演变由个体决定，个体随机地遇到其他个体，个体的看法从“黑”到“白”的转变概率为（$1-\delta$），δ是不进行转换的概率。个体在不遇到其他个体时也可能改变对世界的看法，转变的概率为ε（例如，“外来消息”，或现有交易者被看法不同的交易者所替换，这些原因都可能导致个体改变看法）。

（ⅳ）上述概率按照众所周知的统计过程（马尔科夫链）来进行演变，从k到$k+1$、$k-1$或不发生“变化”的概率分别为

$$k \longrightarrow \begin{cases} k+1 & \text{概率为 } p_1 = p\ (k,\ k+1) \\ k & \text{概率为 } = 1 - p_1 - p_2 \\ k-1 & \text{概率为 } p_2 = p\ (k,\ k-1) \end{cases}$$

其中，

$$p_1 = \left[\frac{1-k}{N}\right]\left[\varepsilon + \frac{(1+\delta)\ k}{N-1}\right]$$

$$p_2 = \frac{k}{N}\left[\varepsilon + \frac{(1-\delta)\ (N-k)}{N-1}\right]$$

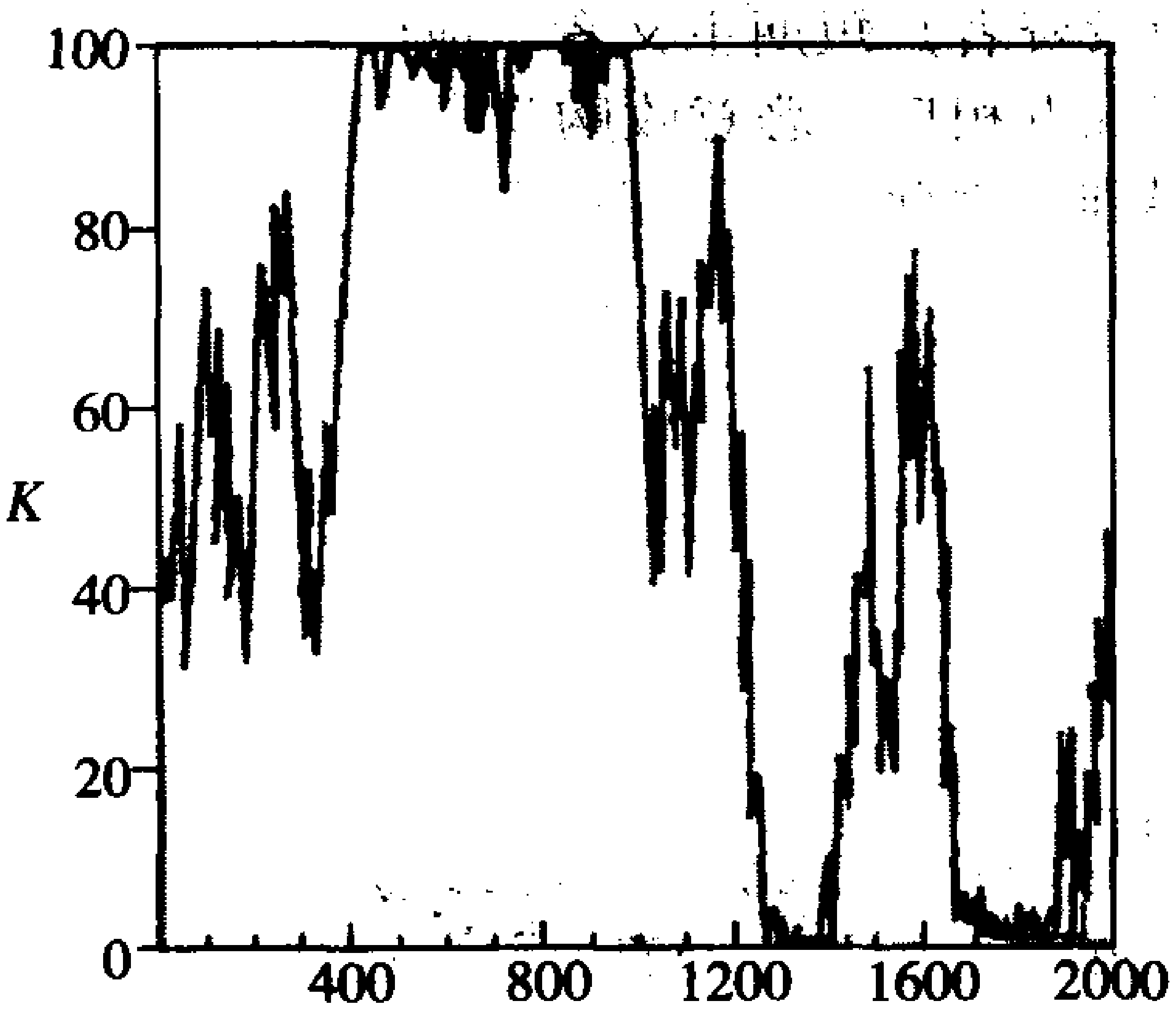


图 2　100 000 次相遇，每 50 次绘一个点：$\varepsilon = 0.002$，$\delta = 0$

资料来源：Kirman（1993a），《经济学季刊》。授权复印。

在 $\varepsilon = \delta = 0$ 的特殊情形中，第一个人总是接受第二个人的观点，动态过程是一个鞅，最后的状态是 $k = 0$ 或 $k = N$。另外，当受他人影响而发生转变的概率（$1-\delta$）较小而自发转变的概率 ε 较高时，两种情形发生的概率都为 50%。Kirman 计算出了系统在两种状态上停留的时间比例（即均衡分布）。结果表明，自发转变的概率 ε 比不转变的概率 δ 越小，系统停留于极端状态的时间就越长，也就是说，100% 的人认为世界是“黑的”或“白的”。（必要条件为 $\varepsilon < (1-\delta)/(N-1)$，参见图 2）。$\delta$ 的绝对值（也就是说，个体“说服力”的强弱）在这里并不重要，重要的是 ε 比 $1-\delta$ 小。因此，个体很有可能转为接受市场中的主流观点，后者是系统演变的主力（也就是说，简单见面使得个体接受主流观点的概率比接受少数人观点的概率高）。

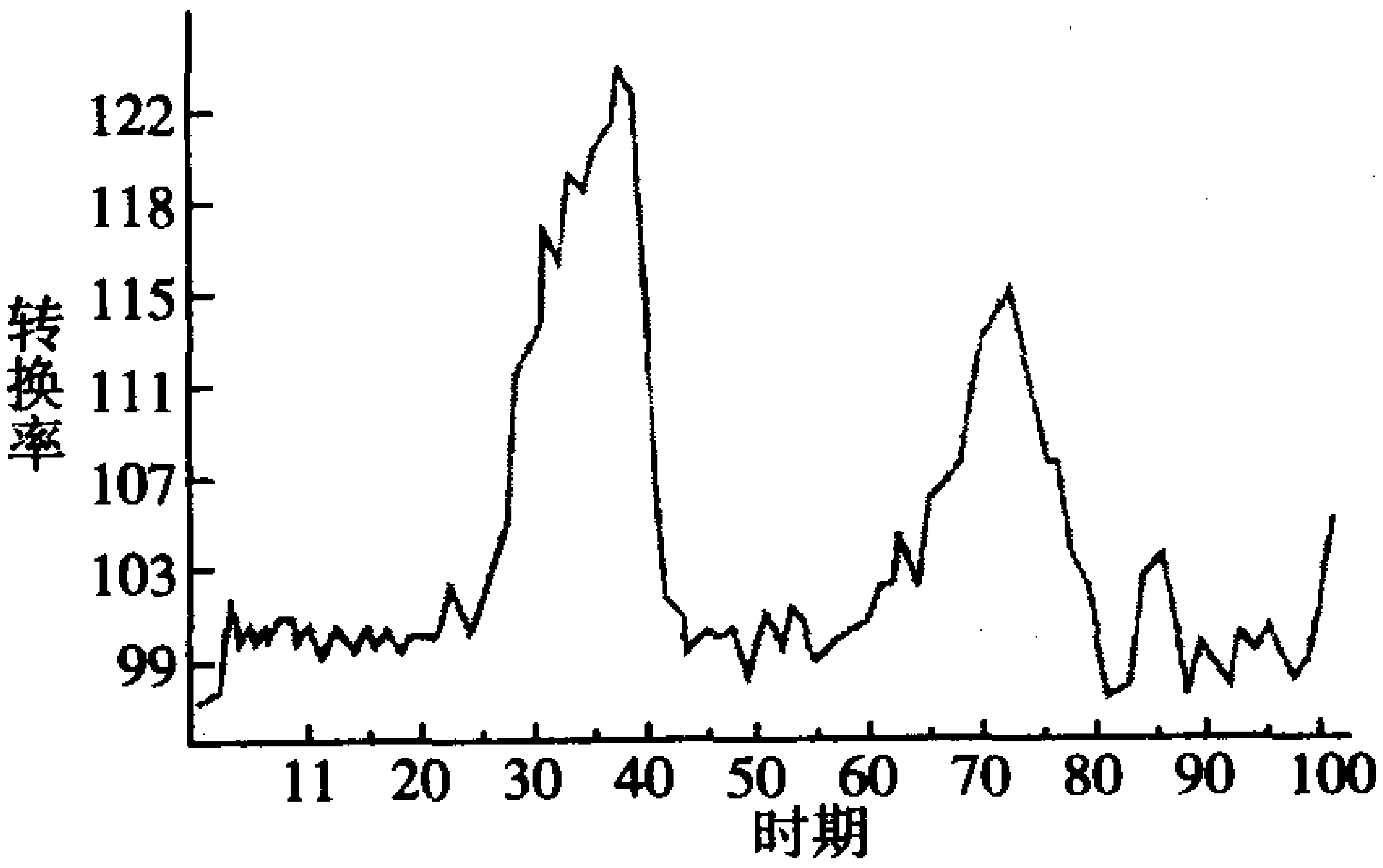


图 3　在 100 期中所拟合的转换率，$S = 100$

资料来源：Kirman（1991），摘自于 Taylor M. P. 主编的《货币与金融市场》，图 17. 3，p. 364。Blackwell 出版社授权复印。

Kirman（1991）使用这类模型来考察汇率等资产可能的行为，资产的价格由基本面交易者看法和噪声交易者看法的加权平均结果来决定。每种类型交易者的比例 w_t 依赖于转换率的上述演变过程，即马尔可夫过程。通过对模型进行模拟，他发现资产（汇率）价格的演变具有一定的周期性，泡沫和衰退交替发生，如图 3 所示。在后来的一篇文章中，Kirman（1993a）

假设基本面价格 p_{ft} 服从随机游走过程，从而该过程是非平稳的，噪声交易者通过简单的外推法来进行价格预测。市场价格的变化 Δp_{t+1} 为

$$\Delta p_{t+1}=w_t\Delta p_{f,t+1}+(1-w_t)\Delta p_{n,t+1} \tag{13}$$

噪声交易者的外推预测为

$$\Delta p_{n,t+1}=p_t-p_{t-1} \tag{14a}$$

基本面交易者的预测

$$\Delta p_{f,t+1}=v(\bar{p}_t-p_t) \tag{14b}$$

是一个误差修正模型，长期均衡值为 $\bar{p}_t$。权重 w_t 依赖于控制市场参与者转换率的参数，权重是内生的，这与 Keynes 的选美冠军思想一致。个体之间相遇后可以转变看法也可以不转变看法，然后他们试图猜测主流的看法，在他们所设想的主流、基本面交易者、噪声交易者看法的基础上来进行预测。所以，代理人进行预测的基础不是他自己的信念，而是她所设想的主流观点。模型模拟结果表明，泡沫和衰退具有周期性。如果噪声交易者完全占主导地位，那么价格为常数；如果基本面交易者占主导地位，那么价格遵循随机游走。单位根标准检验（如 Dickey 和 Fuller（1979），Phillips 和 Perron（1988））适用，p_t 和 $\bar{p}_t$ 之间的协整检验表明，价格泡沫不存在（这是错误的结论）。Hamilton 提出了一种改进的检验方法来检测系统由一个过程转换到另一个过程的时点，但这种方法的成功是有限的。因此，与 Evans（1991）所考察的情形相同，周期性泡沫破灭的时间很难被检测到。

当然，就公共政策含义而言，本章所讨论的所有模型都不能解释一个重要的事实，也就是说，这些模型不能告诉我们，特定股票组合的价格会偏离内在价值多远。例如，如果股票组合的价格与内在价值之间的偏差仅为 5%，即使这种偏离会持续一段时间，在经济中不确定性给定的条件下，这也并不意味着投资基金的配置在很大程度上是错误的。在动荡时期或边际需求持续高涨的特殊时期中，噪声交易者的行为为公共政策提供了一个先例，其形式是交易会停止。噪声交易者的存在也表明恶意收购对收购者来说并不总是很有利，因为他们为目标企业股票所支付的实际价格可能比内在价值高得多。然而，为干预所寻找的表面证据并不能说明政府在市场中的行动就一定是有利的。

19.5　信念与预期

投资者对收益的预测可能存在错误的信念，Barberis、Shleifer 和 Vishny（1998）说明了错误信念是如何导致收益宣告后股票价格的持续反应、短期动量效应、长期均值反转效应以及为什么价格收益比有助于预测未来收益。他们的模型比 Barsky 和 DeLong（1993）的模型更为精致，Barsky 和 DeLong 的模型简单地假设红利增长可以用 IMA（1，1）模型来预测（即适应性预期），这意味着价格变化比当前红利变化大得多（参见第 17 章）。

在公共信息宣告发布后会产生一个系统误差，因为投资者有时会认为收益具有“均值反转”模式，有时甚至会出现一种“趋势”，而收益实际上服从随机游走。投资者必须了解经济运行中的收益模式，并由此来决定他的投资决策。

模型假设投资者属于风险中性类型，经济中存在一个代表性投资者，也就是说，在投资者看到短期正收益后，她会（错误地）应用“小数定律”，并由此认为未来的收益会很高。毕竟收益不会恰好等于“平均收益”，而投资者会错误地认为，即使在小样本情形中，这种平均收益模式在好或坏消息发布后会体现出来。

然而，当我们只有一个收益为正的公告且保守主义者占支配地位时，投资者相信收益会部分地回归正常水平（即均值反转），价格对正面消息具有不足反应。更多的正面消息使得“保守主义者”的人数减少，于是“小数定律”适用，从而收益宣告后股票价格的持续反应和动量效应存在。投资者认为股票处于某种趋势状态中，当收益增长服从随机游走模型时，价格收

益比会上升到比均衡值还要高的位置。当价格相对于内在价值而言很高时，收益增长就不会实现（因为真实收益服从随机游走）。于是，价格具有长期均值反转效应，未来收益会更低，根据历史的价格收益比可以对未来收益进行预测。

意见分歧

Barberis、Shleifer 和 Vishny（1998）的模型考虑了对公共信息的错误认知问题。模型应该怎样考虑私人信息的消化吸收问题呢？假设个体投资者私下会对公司进行研究，所有投资者都持有不同的观点，这些观点都是大家所知道的观点。那些认为股市处于牛市的投资者会购买股票，而那些认为股市处于熊市的投资者会卖空股票。如果不允许后者卖空（如养老基金和大部分共同基金），那么人们会对意见分歧很大的股票存在“更大的乐观性”，这种乐观性在价格中有所体现（Miller（1977））。因此，价格收益比很高，从而后续收益很低，卖出股票者最终会进入市场。由价格收益比所导致的可预测性在这里并不需要出现系统性的牛市行情或熊市行情（与 DeLong、Shleifer、Summers 和 Waldmann（1990a）一样），这里只需要意见不一致就可以了。

对上述结论的检验就需要对意见不一致的程度进行度量。Diether、Malloy 和 Scherbina（2002）使用分析师对当前收益预测的分歧程度来进行度量，然后将股票分为 5 个类别。意见分歧最大类别随后的平均收益比意见分歧最小的类别低。Chen、Hong 和 Stein（2001）使用基金对某只股票的持有比例来度量意见的分歧程度，他们发现分歧较大的股票随后的收益较低。

过于自信

如果投资者对自己关于公司的私人信息过于自信，那么这会产生什么结果呢？如果私人信息是“好”的信息，那么价格会高于内在价值（收益），价格收益比也很高。后来的公共信息会“揭示”这种过高定价，从而未来收益会降低，还可以通过价格收益比来进行预测。Daniel、Hirshleifer 和 Subrahmangan（1998）将过于自信思想与“自我归属偏差”结合起来解释收益宣告后股票价格的持续反应和动量效应。投资者对待公共信息的态度不对称。如果在对正面私人信息产生过于自信之后紧随的是正面的公共信息，那么投资者会将正面的公共信息“归因于”自己的技能，他会变得更加自信，并使价格进一步上升。净效应是更加过于自信，也会产生动量效应。在 Barberis、Shleifer 和 Vishny（1998）的“错误预期”模型和上述“过于自信”模型中，对股票价格的绝大多数修正应出现在投资者发现他们的私人信息的确是错误的时候（也就是说，出现在收益公告日）。这与 Chopra、Lakonishok 和 Ritter（1992）及 La Porta、Lakonishok 和 Shleifer（1997）的实证结果一致，他们发现，有些股票组合（如价值型或增长型股票）在接近收益宣告时所获得的平均收益较低。

19.6　动量与信息观察者

在 Hong 和 Stein（1999）的有趣模型中，同质交易者之间存在交互作用。动量交易者（趋势交易者）的投资决策仅仅基于历史价格的变化，而“信息观察者”只能观察到私人信息，这些私人信息在信息观察者之间逐渐扩散。两类投资者都是有限理性的，因为他们不使用所有可获得的信息。如果只有信息观察者，那么价格反应就会千篇一律，不足反应也会存在，因为消息被价格消化吸收的速度非常缓慢，一旦引入动量交易者，过度反应也就不存在。在 t 时刻出现好消息之后，由于信息观察者需求的增加，价格会上涨，这会导致一些动量交易者在 $t+1$ 时刻的需求出现增加，这又会导致价格出现加速上涨，动量交易者也会疯狂购买。在“动量周期”的早期阶段，动量交易者能获得绝大部分利润。买得比较晚的动量交易者（在 $t+i(i>1)$ 时刻购买）会遭受损失，因为价格已经大大高于长期均衡价格，因此，一些动量交易者在价

格到达顶峰时才开始购买（也就是说，“晚”进入的动量交易者会受到负的外部性的影响）。动态过程是市场出清时的均衡模型结果，但有限理性假设（如动量交易者只能使用单变量预测，而不知道新信息在什么时候到达）对短期正自相关性和长期价格反转而言都是非常关键的。

如果我们在模型中引入拥有完全信息的“理性套利者”，只要理性套利者的风险承受能力有限，那么上述结论仍然成立（如果理性套利者的风险承受能力是无限的，那么价格服从随机游走。）

因此，模型的预测结果与所观察到的动量交易者获得利润的现象一致（Jegadeesh 和Titman（1993，2001）），可能的原因是，私人信息在开始的时候扩散得非常缓慢。事件研究文献表明，所观测到的公共事件（如预期不到的良好收益、新股发行、股票回购、分析师的推荐）会导致事件后6~12个月的时期中出现股票价格的持续反应（与初始事件的方向一致）。Hong-Stein 的模型只有在如下条件下才能产生股票价格的持续反应效应：当观察到公共信息时，在信息观察者将公共信息融入到对未来价值的看法之前，他们需要一些额外的私人信息。否则，信息观察者有可能立即将公共信息与股票价格联系起来。

模型的基本结构相当复杂，我们在这里只强调模型的主要特征。信息观察者购买风险资产，这种资产在 T 时刻存在红利支付 D_T，

$$D_T = D_0 + \sum_{j=0}^{T} \varepsilon_j$$

$\varepsilon_j \sim iid\ (0,\ \sigma^2)$。为了使私人信息通过信息观察者群体进行缓慢扩散，我们假设有 z 组信息观察者，从而红利创新（innovation）为 $\varepsilon_i = \varepsilon_j^1 + \varepsilon_j^2 + \cdots + \varepsilon_j^z$（每个子创新的方差为 σ^2/z）。正如我们即将看到的那样，如果存在 z 组信息观察者，那么 t 时刻的红利创新要到 $t+z-1$ 时刻才为每一个信息观察者所吸收。因此，私人信息在 $t+z-1$ 期之后才会变成公共信息。我们使用如下程序来进行模拟。在 t 时刻，第1组信息观察者观测到 ε_{t+z-1}^1，第2组信息观察者观测到 ε_{t+z-1}^2，以此类推，每一组都观测到创新 ε_j 的 $1/z$。于是，在 $t+1$ 时刻，第1组观测到 ε_{t+z-1}^2，第2组观测到 ε_{t+z-1}^3，其余以此类推，第 z 组现在观察到的是 ε_{t+z-1}^z。在 $t+1$ 时刻，信息观察者能观察每一个子创新 ε_{t+z+1} 的 $2/z$，信息进一步扩散。所以，ε_{t+z+1} 在 $t+z-1$ 时刻完全变成了公共信息，但信息的扩散很缓慢（扩散速度取决于信息观察者的数目 z，它是信息扩散速度一个度量指标）。一般而言，所有信息观察者获得的信息量都相等。

每个信息观察者的绝对风险规避系数为常数（CARA $= \gamma$），他们在 t 时刻的资产需求来源于静态最优化（即买而持有，一直持有到 T 时刻）。最为重要的是，他们不会根据历史价格来进行决策（因此，他们是有限理性的）。在没有损失的情形中，无风险利率为0，信息观察者的期限为无穷大。

动量交易者也可能具有 CARA 效用函数，但期限是有限的。在每个时刻 t，新的动量交易者群体进入市场，持有头寸一直到 $t+j$ 时刻。动量交易者在 $P_{t+j} - P_t$ 预测值的基础上确定资产需求，只根据 ΔP_{t-1} 来进行预测（即单变量预测）。如果允许动量交易者根据历史价格的滞后分布（权重不同）来预测价格，那么模型的结果不成立，这就是有限理性假设。t 时刻动量交易者的需求指令流 F_t 为

$$F_t = A + \phi \Delta P_{t-1}$$

t 时刻有 j 个动量交易者，假设他们的需求被信息观察者（扮演了造市商的角色）所消化和吸收。因此，如果风险资产的供给是固定的且等于 Q，那么信息观察者就必须吸收 S_t，

$$S_t = Q - \sum_{i=1}^{j} F_{t+1-i} = Q - jA - \sum_{i=1}^{j} \phi \Delta P_{t-i}$$

可以证明，均衡价格（使市场出清）为

$$\begin{aligned} P_t = {} & [D_t + \{ (z-1)\ \varepsilon_{t+1} + (z-2)\ \varepsilon_{t+1} + \cdots + \varepsilon_{t+z-1} \} / z] \\ & - Q + jA + \sum_{i=1}^{j} \phi \Delta P_{t-i} \end{aligned}$$

可以证明，均衡状态下的 ϕ 取决于动量交易者的风险忍耐系数 $1/\gamma$。中括号内的项意味着价格存在自相关性。模型要用数值方法进行求解，敏感性分析给出了如下定性结果：

（ⅰ）对持有期为 $j=12$ 个月而言，过度反应达到最大值，约为34%。

（ⅱ）随着风险忍耐力的增加，动量交易者对历史价格变化的反应更为强烈，均衡值 ϕ 增加，过度反应更为剧烈。

（ⅲ）随着信息扩散参数 z 的增加（从而私人信息扩散的速度“更慢”），动量交易者更具有进取性（即 ϕ 增加），他们的利润也会增加。过于反应也会增加，这使得价格在反转时期具有更大的负自相关性。

（ⅰ）中的结果与 Jegadeesh 和 Titman（1993，2001）的发现一致，动量交易利润在12个月左右发生。最后一个结果（ⅲ）为模型提供了进一步检验的方法，前提是我们能将那些信息扩散相对较慢的股票分离开来。Hong、Lim 和 Stein（2000）使用“公司规模”和“分析师报告”（即对公司规模进行修正后的报告）来刻画信息缓慢传播的速度。他们发现，6个月动量交易利润随着市场规模的增加而减小，随着分析师报告数量的增加而增加。此外，对分析师报告比较少的股票而言，动量交易利润会持续2年左右；对分析师报告比较多的股票而言，动量交易利润的持续时间小于一年。

上述模型的假设个数已经达到最小，但模型存在一些局限。信息观察者具有时间不一致性，他们在 t 时刻在静态框架下来决定资产需求。但是，他们在观察到动量交易者需求之后也会改变自己的需求。信息观察者不能成为领先者。如果信息观察者的需求依赖于历史数据，那么他们清楚好消息能增加动量交易者的需求，所以在时刻买入更多的股票，从而能够从动量需求预测中获利。这可能会减少一些不足反应现象（但不能完全消除这种弱式反应）。

模型对私人信息反应不足，但这并不意味着模型对公共信息（比如收益公告）也反应不足。然而，如果在公共信息吸收方面花费的时间等于私人代理人承担对“公共信息给价格带来的影响”进行预测所花费的时间，那么模型能反映出事后持续效应。当然，如果动量交易者能观察到这些公共信息，他们会进行快速交易，最终不存在对公共信息的过度反应，因此价格反转也就不存在。然而，必须注意的是，代表性投资者模型不能得出上述（ⅰ）～（ⅲ）类型的结果。它们仅仅依赖于立即就可以得到的信息。

19.7 风格投资

在上一章中，我们提到了风格投资（style investing）的盛行。漫不经心的经验主义者认为，许多共同基金都有一定的投资风格（如价值增长、动量、小市值、高科技股、房地产等）。可能出现的情形是，投资者在几种风格上配置财富，而不很关心某种风格内各种股票之间的财富配置问题。与个体股票分析基础上的投资组合构造相比，风格投资肯定存在一定的监管成本和和交易成本（Bernstein（1995），Swensen（2000））。我们已经遇到过一些依赖于动量交易者与某种形式的理性交易者或套利交易者间交互作用的模型，如 DeLong、Shleifer、Summers 和 Waldmann（1990a）、Hong 和 Stein（1999）。在这些模型中，动量（噪声）交易者的股票需求依赖于股票的绝对历史绩效。Barberis 和 Shleifer（2003）的风格投资模型假设动量交易者对特定投资风格 X 中股票的需求取决于 X 相对于另外一种风格 Y 的历史收益（例如，X = 旧经济中的股票，Y = 新经济中的股票）。因此，如果 X 的历史收益大于 Y，那么动量策略投资者就会由 Y 风格转向 X 风格，这会增加 X 风格资产的收益，降低 Y 风格资产的收益。两种不同风格资产的收益在滞后几期后存在一种负的相关性。但 Y 风格的资产最终会回归内在价值，因此，在滞后多期之后，$\Delta P_{X,t}$ 和 $\Delta P_{Y,t-k}$（k 的值很大）之间存在正的相关性。

模型中存在套利交易者或基本面交易者，他们的需求依赖于基于基本面的预期收益估计值。套利交易者与造市商一样扮演了“转换者”的角色，吸收需求变化方面的信息。模型给

出了所有资产的市场出清价格，但是，因为基本面交易者是有限理性的，不知道（转换）需求变化的时间序列性质，所以市场出清价格与只存在基本面交易者时的价格不同。

根据上述描述我们可以看出，一般均衡模型的宽泛设定与 Hong 和 Stein（1999）有着共同相同的特征，但对“转换者”和噪声交易者行为的不同假设确实导致了对股票价格的不同预测。在考察后一结论之前，我们简要地陈述模型的主要部分，导出 X 风格资产和 Y 风格资产的市场均衡价格。

假设存在 $2n$（比如说为 100）种风险资产（无风险资产具有无穷大的供给弹性，收益为 0），它们的供给是给定的。每一种风险资产都是对未来某个时刻支付的红利拥有索取权利：

$$D_{i,T}=D_{i,0}+\varepsilon_{i,1}+\varepsilon_{i,2}+\cdots+\varepsilon_{i,T}$$

ε_{it}在 t 时刻到达，$\varepsilon_t=(\varepsilon_{1,t},\ \varepsilon_{2,t},\cdots\ \varepsilon_{2n,t})'\sim N(0,\ \Sigma_D)$，$\varepsilon_t\sim iid$。经济中存在 X 和 Y 两种风格的资产，资产 1 到 n 属于 X 风格，资产 $n+1$ 到 $2n$ 属于风格 Y。X 风格资产的收益为

$$\Delta P_{X,t}=P_{X,t}-P_{X,t-1}$$

$$P_{X,t}=n^{-1}\sum_{i\in X}P_{i,t}$$

每一个资产冲击 ε_{it}都依赖于市场因子（$=m$，对两种风格而言都相同）、风格因子（$s=X$ 或 Y）和单一资产所特有的异质型现金流冲击。这些因子的方差都为 1，每一个因子与其他因子都是正交的，因此

$$\Sigma_D^{ij}=\mathrm{Cov}(\varepsilon_{it},\varepsilon_{jt})=\left\{\begin{array}{ll}1 & i=j\\ \psi_m^2+\psi_s^2 & i,j\text{ 属于同一种风格,但 } i\neq j\\ \psi_m^2 & i,j\text{ 属于不同的风格}\end{array}\right\}$$

转换者的需求取决于两种风格资产的相对历史绩效。对 X 风格中的资产 i 而言，需求数量为

$$N_{i,t}^s=\frac{1}{n}\sum_{k=1}^{t-1}\theta^{k-1}\ (\frac{\Delta P_{X,t-k}-\Delta P_{Y,t-k}}{2})\ =\frac{N_{X,t}^s}{n}$$

$0<\theta<1$，这意味着我们对历史上的相对收益赋予递减的权重。根据对称性，对 Y 风格中的资产 j 而言，

$$N_{j,t}^s=\frac{1}{n}\sum_{k=1}^{t-1}\theta^{k-1}\ (\frac{\Delta P_{Y,t-k}-\Delta P_{X,t-k}}{2})\ =\frac{N_{Y,t}^s}{n}$$

上述命题假设，当风格投资者希望转换投资风格时，他们在同一资产类别中（如股票）将需求从一种风格转移到另外一种风格（如从价值型股票转移到成长型股票），而不是将需求从一种资产转移到另外一种资产（如现金、债券和外汇之间的转换）。这对机构投资者而言大体是真实的，机构投资者在各个资产类别上有着非常稳定的资产配置“策略”。交易成本也意味着，广受欢迎的投资风格是通过售卖一种（或几种）绩效差的风格资产来进行融资，而不是售卖许多种风格资产来进行融资。因为一些“经验规则”导致了孪生风格（twin styles）（如价值型与增长型）的存在，所以当一种风格的资产表现很好时，“孪生风格”资产总是表现得很糟糕。

基本面交易者追求预期效用的最大化，CARA 效用函数是期末财富的函数。因此，当收益服从正态分布时，最优资产需求是预期收益的线性函数

$$N_t^F=\ (1/\gamma)\ (V_t^F)^{-1}\ [E_t^F\ (P_{t+1})\ -P_t]$$

其中，$V_t^F\equiv\mathrm{Var}_t^F\ (P_{t+1}-P_t)$，$N_f^F=\ (N_1,\ N_2,\ \cdots,\ N_{2n})'$，$P_t=\ (P_1,\ P_2,\ \cdots,\ P_{2n})'$，$\gamma$ 是绝对风险规避系数。如果 $2n$ 种资产的固定供给由向量 $Q=N_t^F+N_t^S$ 来确定，那么将 $N_t^F=Q-N_t^S$ 代入上述方程就得到

$$P_t=E_t^F\ (P_{t+1})\ -\gamma V_t^F\ (Q-N_t^S)$$

向前迭代并注意到 $E_{T-1}^F P_T=E_{T-1}^F\ (D_T)\ =D_{T-1}$就得到

$$P_t=D_t-\gamma V_t^F\ (Q-N_t^S)\ -E_t^F\sum_{k=1}^{T-t-1}\gamma V_{t+k}^F\ (Q-N_{t+k}^S)$$

其中，$D_t=(D_{1,t},D_{2,t},\cdots,D_{2n,t})'$。假设协方差矩阵 V^F 具有时间不变性（$=V$），且与现金流协方差矩阵 Σ_D 具有相同的结构。假设 $E_t^F(N_{t+k}^S)$ 为常数，从而基本面交易者是有限理性的，他们不会计算 N_{t+k}^S 的时间序列性质，只是吸收转换者需求的一些信息。扔掉所有的非随机项就得到

$$P_t=D_t+\gamma VN_t^S$$

X 风格中资产 i 的价格为

$$P_{it}=D_{it}+\left(\frac{\psi}{n}\right)\sum_{k=1}^{t-1}\theta^{k-1}\left(\frac{\Delta P_{X,t-k}-\Delta P_{Y,t-k}}{2}\right)$$

其中，ψ 直接依赖于 γ 和 Σ_D（$=V$）的协方差结构参数。风格 Y 中资产 j 的价格与上式有着相同的形式，但（$\Delta P_X-\Delta P_Y$）的符号反过来了。如果只有基本面交易者，那么 $P_t=D_t$。如果存在转换者，那么价格会偏离内在价值。如果 θ 接近 1，那么这种偏离会一直持续下去。

模型可以用 $\psi_m=0.25$、$\psi_s=0.5$、$\theta=0.95$ 和 $\lambda=0.093$ 来进行校准（均衡收益的波动率基本上与美国数据一致）。这反过来又得到 $\phi=(\psi/n)^{-1}=1.25$。他们取 $n=50$，从而每种风格有 50 种资产，对所有的 i 而言，$D_{i0}=50$。

预测与协同运动

首先讨论风格收益（style returns）的协同运动。如果 X 风格存在一个一次性的现金流冲击（即对 $t>1$ 和 $\forall i\in X$ 有 $\varepsilon_{i,1}=1$，$\varepsilon_{i,t}=0$），那么 P_X 会围绕（新的更高的）内在价值进行长期变动，在初始时刻，P_X 高于长期均衡价格，随后缓慢地出现均值反转现象。其他动量模型（如 Hong 和 Stein（1999），DeLong、Shleifer、Summers 和 Waldmann（1990a））也能预测到这种短期正自相关性和长期负自相关性。原因很直观。X 风格资产现金流方面的利好消息会导致价格上升，从而增加转换者的需求，推高价格并使得价格高于内在价值。这种风格“方法”的一个新特征是，Y 风格资产的价格在开始时会下降，因为 X 风格资产的购买要通过卖出 Y 风格资产来进行融资（即对称效应）。相对于 X 风格资产的收益而言，这使得 Y 风格资产看上去非常糟糕，因此，对 Y 风格资产而言，动量售卖效应会增加，Y 风格资产的价格最终会低于内在价值。注意，Y 风格资产的价格运动是在没有任何现金流方面消息的情形下进行的，不同风格资产收益之间的相关性在短期为负。最终的结果是，基本面交易者卖出价格高估的 X 风格资产，价格向长期均衡水平回落。X 风格资产的坏消息或 Y 风格资产的好消息都会使这一过程加速。

因此，X 风格对 Y 风格资产施加了一种外部性，外部性的大小取决于投资者如何对 X 风格资产的购买进行融资。与投资者只卖出孪生风格中的 Y 风格资产相比，如果他们卖出少量的多种不同风格的资产，那么这种外部性会更小一些。

由上述分析可以看出，如果（比方说）X 风格中只有一种资产经历了一次性正的现金流冲击，那么 X 风格中其他资产的价格也会出现上涨（上涨与现金流无关），而 Y 风格中资产的价格则会下跌（下跌同样与现金流无关）。在 Hong 和 Stein（1999）及 DeLong、Shleifer、Summers 和 Waldmann（1990a）的模型中，资产 i（$i\in X$）的利好消息只影响资产的价格，而不会影响 X 风格中其他资产的价格，也不会影响 Y 风格资产的价格。从这一点来看，这些模型的预测结果不同。

上述结果与 1979—1983 年期间人们在小市值公司上的成功投资一致（Banz（1981）），1983 年后小市值公司股票收益变得非常糟糕。数据挖掘（data snooping）也会导致与现金流消息无关的收益变化，而某些风格（如美国 1998 和 1999 年的价值股）资产的糟糕收益（即使现金流方面的消息是利好消息（Chan、Karceski 和 Lakonishok（2000）））应归因于替代风格股票（如大的成长型股票）需求的增加。

如果某种风格内的资产收益在历史上处于相对较好的状况，那么投资者会在这种风格中的所有证券上进行移动。因此，这一特殊风格内的资产收益之间可能存在正的协同运动，且运

动与共同的现金流资源无关。这与封闭式基金价格的协同运动现象一致，即使这些基金的净资产价值之间只是微弱相关（Lee、Shleifer 和 Thaler（1991））。这是 DeLong、Shleifer、Summers 和 Waldmann（1990a）及 Hong 和 Stein（1999）的模型所无法预测的，在他们的模型中，个体资产收益由潜在的现金流来决定；这种协同运动也不是学习模型所能预测的（Veronesi（1999），Lewellen 和 Shanken（2002））。当然，其他模型也能解释与现金流无关的某些协同运动。例如，在 Kyle 和 Xiong（2001）中，在银行遭受交易损失后，他们会卖出股票来弥补资本金。这与1998年的金融危机相一致，各个国家的总体股票价格大幅下跌（即使不同国家的经济基本面不同），但它不能很好地解释某些股票子类（如小市值股票）的协同运动现象。

依据基本面分析方法，因为皇家荷兰公司与壳牌运输公司的股票有着相同的现金流，所以他们的价格应该非常接近。但是，Froot 和 Dobroa（1999）认为，皇家荷兰公司的股票价格与美国市场同步变动，而壳牌运输公司的股票价格与英国市场同步变动。皇家荷兰公司的股票主要在美国进行交易，而壳牌运输公司的股票主要在英国进行交易，它们属于两种风格的资产。类似地，如果某支股票被编入 S&P500 指数（即一种风格），那么我们可以预期它在将来会与 S&P500 指数协同运动，它与 S&P500 外的其他股票之间的相关性会下降（Barberis 和 Wurgler（2001））。

最后，Barberis 和 Shleifer 认为，基于风格的投资策略（也就是说购买风格内绩效较好的股票）的夏普比至少与基于个体资产动量绩效的动量策略夏普比一样大。这与基于行业组合的动量策略（Moskowitz 和 Grinblatt（1999））和根据规模及账面市值比对动量组合进行排序的动量策略（Lewellen（2002））一致。

即使在模型中引入更加聪明的套利交易者（他们明白动量交易者需求的时间变化规律），这也不必然减小错误定价的大小和持续性。这是因为，与 DeLong、Shleifer、Summers 和 Waldmann（1990a）的模型一样，套利交易者不会在价格高于内在价值时卖出资产（他们会买入资产），因为他们知道反馈（feedback）交易者逐渐增加的需求将进一步提升资产价格，然后套利交易者在获利后退出市场。换句话说，套利交易者在股价上升（或下降）后会模仿动量交易者的行为。

19.8 前景理论

前景理论（prospect theory）是以 Kahneman 和 Tversky（1979）心理学实验为基础的风险决策描述模型。他们注意到，个体经常参加与期望效用理论中公理相违背的一些赌博。他们的实验意味着，个体关心的是财富变化（而不是财富的绝对水平），与收益相比，个体对损失更加敏感。这即使著名的损失规避（loss aversion）。然而，要注意的是，尽管前景理论与只取决于消费的标准预期效用函数相去甚远，人们还是可以使用其他形式的目标函数，但这些目标函数不依赖于预期效用。这些模型包括失望规避（disappointment aversion）（Gul（1991））、使用收益的各种分布来追求最小预期效用（minimum expected utility）的最大化（Gilboa 和 Schmeidle（1989）、Camerer 和 Weber（1992））、使用加权后的效用函数（Chew（1989））或与排序有关的效用函数（Segal（1989））/使用稳健性规则（Anderson、Hansen 和 Sargent（2000））。

下面的模型体现了前景理论的思想，模型中不存在与理性套利者相互影响的外生噪声交易者。相反，它完全是一种最优化方法，但代理人的效用函数不是标准效用函数：一生的效用不仅取决于消费，还取决于风险资产最近的收益和损失（即股票市场上财富的变化）。这是一个狭隘构架（narrow framing）假设，因为模型忽略了其他资产的收益或损失，即使投资的期限很长，投资者也会担心每年的收益和损失。损失会影响人们的资产配置这一思想已经得到试验证据的支持。Thaler、Tversky、Kahneman 和 Schwartz（1997）将股票和债券的收益分别告诉三组投资者（为了模拟真实收益，股票和债券的收益根据校准后的两个正态分布来生成）：第一

组看到的是月度观察值，第二组看到的是年度观察值，第三组看到的是每五年一次的观察值（时间不重叠）。在每一种期限的末端，投资者都要将资金配置到股票和债券中去，还要报告他们的收益或损失。经过200个月的观测后，每一组都要求进行最后一次资产配置，配置的期限是接下来的400个月（约33年）。结果发现，第一组比第二组和第三组的损失更大，第一组在最后的配置中将更少的资金投资到股票上。

Benartzi 和 Thaler（1995）的单期模型证明了损失规避可能导致很高的股权溢价，但 Barberis、Ming 和 Tano（2001）的跨期模型却发现他们需要一种附加效应：某种类型的积分控制，他们称之为事前损失（prior losses）（Thlaer 和 Johnson（1990））。这种思想很直接、直观，也很合理，也就是说，如果你在前几期中遭受了损失，那么这一期所遭受的损失会使你感到特别痛苦，从而你更加厌恶风险。反过来的情形也一样，在获得一系列的收益之后（如卡西诺赌场），个体更愿意参与赌博，因为他们现在“玩的是庄家的钱”，也就是说，他们玩的是卡西诺赌场的钱而不是他们自己的钱。Barberis、Ming 和 Tano（2001）证明了如下结论：损失规避加上“事前损失”可以解释一些典型事实，如高股权溢价、股票收益的高波动性、利率较小的水平值和较小的波动率、股票收益的可预测性（如低的价格红利比可能导致更高的未来收益）。

显然，时变风险规避的思想与 Campbell - Cochrane 习惯持续模型相似，但是，导致风险规避程度发生变化的是财富的累积变化而不是历史消费的累积水平。正如我们即将看到的那样，Barberis 等使用校准方法来证明，如果参数取“合理”值，那么模型就能得到与典型事实一致的股票收益和利率时间序列行为。

经验证据表明，投资者倾向于卖出获利的股票而不是输钱的股票（Shefrin 和 Statman（1985）、Odean（1998）），尽管这些研究假设效用收益和损失只在卖出股票后才会发生，这些经验证据基本上与事前损失和损失规避一致。在 Barberis 等的模型中，即使没有资产收益和损失，效用也存在收益或损失。

模型

市场上有两种资产，无风险资产净的供给为0，毛利率为 $R_{f,t}^{*}$，风险资产的收益为 R_{t+1}^{*}（在 t 到 $t+1$ 的时期上）。风险资产的总供给为1个单位，拥有对红利序列 $\{D_t\}$ 的索取权。经济中有无穷多个代理人，每人在 $t=0$ 时刻拥有1单位的风险资产，他们在任何时刻都持有风险资产。在“经济 I”中，代理人消费红利流（Lucas（1978）），总消费 C_t 等于总红利 D_t，因此两者服从相同的随机过程（即独立同分布，且分布为对数正态分布），

$$c_{t+1} - c_t = g_c + \sigma_c \varepsilon_{t+1} \quad \varepsilon_{t+1} \sim niid\ (0,\ 1) \tag{15}$$

在“经济 *II*”中，消费和红利服从不同的随机过程，但他们各自都是独立同分布的，但均值和标准差不同，它们的残差之间具有非零的相关性（参见下文）。

效用函数是关于消费的幂函数，函数 $v(.)$ 表示在 t 到 $t+1$ 的时期中所获得的收益或所遭受的损失 X_{t+1}，

$$U = E_t\left[\sum_{t=0}^{\infty} \theta^t \frac{C_t^{1-\gamma}}{1-\gamma} + b_t \theta^{t+1} v\ (X_{t+1},\ z_t)\right] \tag{16}$$

γ（>0）是关于消费的相对风险规避系数，$v(X_{t+1},\ z_t)$ 是由收益或损失所导致的效用，z_t 代表事前收益或损失。如果没有事前收益或损失，那么 $z_t = 1$（下面会进行更为深入的解释）。b_t 是外生的比例因子。所选择的期限单位为年。如果 $S_t = 100$ 美元是参考水平，那么

$$X_{t+1} = S_t\ (R_{t+1}^{*} - R_{f,t}^{*}) \tag{17}$$

如果 $R_{t+1}^{*} = 1.1$ 且 $R_f^{*} = 1.05$，那么 $X_{t+1} = 15$ 美元。Kahneman 和 Tversky（1979）的损失规避效用函数是分段线性函数：

$$v\ (X_{t+1},\ z_t = 1) = \begin{cases} X_{t+1} & X_{t+1} \geqslant 0 \\ \lambda X_{t+1} & X_{t+1} \leqslant 0 \end{cases} \tag{18}$$

$\lambda > 1$ 是风险规避的反映，这里没有事前收益或损失。这个效用函数与小额财富赌博中所观察到的风险规避一致。与小额赌博相匹配的普通光滑效用函数（如幂效用函数），对大额赌博而言，会给出非常荒谬的结果。例如，Rabin（2000）表明，那些拒绝参加以 50∶50 的概率损失 100 美元和获得 110 美元赌博的效用最大化者，也会拒绝参加那些以 50∶50 的概率损失 1000 美元和获得任意数量（包括无穷大）美元的赌博。损失规避者会回避“Rabin 悖论”。

在 Barberis 等的模型中，重要的是期望效用 $v(X_{t+1})$，他们假设收益和损失的主观概率相等（与此不同的是，Kahneman 和 Tversky（1979）中的概率是这些概率的非线性变换）。由（17）和（18）可知，效用函数依赖于收益

$$v(X_{t+1}, z_t = 1) = v[S_t(R^*_{t+1} - R^*_{f,t})] = S_t\hat{v}(R^*_{t+1}) \tag{19}$$

其中，

$$\hat{v}(R^*_{t+1}, z_t = 1) = \begin{cases} R^*_{t+1} - R^*_{f,t} & R^*_{t+1} \geqslant R^*_{f,t} \\ \lambda(R^*_{t+1} - R^*_{f,t}) & R^*_{t+1} < R^*_{f,t} \end{cases} \tag{20}$$

（16）中的比例因子 b_t 是为了使财富 S_t 是时间的递增函数，但它不会在效用中占主导地位。我们可以把总消费或财富用作比例因子，Barberis 等选择了前者（对个体投资者而言它是外生的）。

$$b_t = b_0\bar{C}_t^{-\gamma} \quad (b_0 > 0) \tag{21}$$

如果 $b_0 = 0$，那么（16）中的效用函数就变成我们熟悉的幂效用函数，幂效用函数只是关于消费的函数，如果越大，那么效用函数就会给财富变化（即收益）而不是消费赋予更大的权重。

事前结果

在获得一系列收益（遭受损失）后代理人的风险规避程度会降低（增加），这一思想来自于对如下问题的调查（Thlaer 和 Johnson（1990））：

1. 你刚赢了 30 美元。请选择：

（a）50% 的机会得到 9 美元，50% 的机会损失 9 美元［8%］

（b）没有其他损失或收益［1%］

2. 你刚刚损失了 30 美元。请选择：

（a）50% 的机会得到 9 美元，50% 的机会损失 9 美元［3%］

（b）没有其他的损失或收益［6%］

括号中的百分比是进行相应选择的人数比例，因此，在已经损失 30 美元后，人们就更不愿意参加赌博。

Gertner（1993）表明，这种“玩庄家钱”的效应在更大的赌博中也会出现，在 TV 赌博秀中，参与者对下一次随机抽取的卡片比当前卡片大还是小进行下注。Linville 和 Fisher 使用调查证据后发现，投资者希望不愉快的事情间断发生而不是一起发生，他们也希望在“好”事情之后紧接着出现“坏”事情，从而好坏事情之间有一个缓冲作用。

为了描述事前结果的概念，我们需要风险资产的历史参考水平 Z_t，$S_t - Z_t$ 反应的是财富的变化。当 $S_t > Z_t$ 时，投资者的风险规避程度比平常低。测度事前结果的状态变量为 $z_t = Z_t/S_t$，$z_t < 1$ 表示存在可观的事前收益，因此风险规避程度要低一些。效用由 X_{t+1} 和 z_t 来决定，因此我们有 $v(X_{t+1}, z_t)$。$v(X_{t+1}, z_t)$ 的函数形式很简单，但很精巧，它考虑了如下两种情形：（a）效用损失取决于事前收益或损失的大小，（b）事前收益比事前损失对效用函数的影响要大一些。为了简便起见，我们设 $R^*_f = 1$（对 $R^*_f \neq 1$ 的情形，可参见 Barberis、Ming 和 Tano（2001）），根据基准水平将收益分成两个部分。

假设开始的基准水平为 $Z_t = 90$ 美元，$S_t = 100$ 美元，从而 $z_t = 0.9$，即事前收益存在。如果股票价格在 $t+1$ 期下降到 $S_tR^*_{t+1} = 80$ 美元，那么我们不会对所有的损失（20 美元）进行惩罚，即 $\lambda = 2.0$（比方说）。从 100 美元到基准水平 90 美元的损失由权重设 1 来进行惩罚，低于基准水平的损失（即从 90 美元到 80 美元的损失）由权重 $\lambda = 2.0$ 来进行惩罚。因此，20 美元损

失所带来的负面效用为

$$效用变化 = (90-100)(1) + (80-90)(2.0) = -30 \tag{22}$$

$$\begin{aligned}效用变化 &= (Z_t - S_t)(1) + (S_t R^*_{t+1} - Z_t)\lambda \\ &= S_t(z_t - 1)(1) + S_t(R^*_{t+1} - z_t)\lambda\end{aligned} \tag{23}$$

如果损失足够的小，$S_t R^*_{t+1} > Z_t$，或等价的说 $R^*_{t+1} > z_t$，那么所有损失都由权重 1 来进行惩罚。我们分两种情形来考虑对效用的影响：事前损失和事前收益。

情形 A：事前收益，$z_t \leqslant 1$

$$v(X_{t+1}, z_t) = \begin{cases} S_t(R^*_{t+1} - 1) & R^*_{t+1} \geqslant z_t \\ S_t(z_t - 1) + \lambda S_t(R^*_{t+1} - z_t) & R^*_{t+1} < z_t \end{cases} \tag{24}$$

因此，如果 $S_t = 100$ 美元，$Z_t = 90$ 美元，$S_t R^*_{t+1} = 90$ 美元，那么 $R^*_{t+1} = z_t = 0.9$，当前收益 $R^*_{t+1} - 1 = -10\%$ 刚好抵消 10% 的历史事前收益，效用损失的权重为 1。如果 $S_t R^*_{t+1} = 80$ 美元，那么 $R^*_{t+1} = 0.8 < z_t = 0.9$，当前损失为 20%，这意味着 $S_t R^*_{t+1} = 80$ 美元低于历史参考水平 $Z_t = 90$，效用损失导致"痛苦"更大一些（即（24）中的 λ 要起作用）。

情形 B：事前损失，$z_t > 1$

相似的情形适用于事前损失，除非我们假设更多的损失会导致更大的痛苦，从而 λ 是事前损失函数的递增函数，

$$\lambda(z_t) = \lambda + k(z_{t-1}) \qquad (z_t > 1 \text{ 且 } k > 0) \tag{25}$$

且

$$v(X_{t+1}, z_t) = \begin{cases} X_{t+1} & X_{t+1} \geqslant 0 \\ \lambda(z_t)\, X_{t+1} & X_{t+1} < 0 \end{cases} \tag{26}$$

令 $\lambda = 2$，假设 k 的取值为 10。令 $Z_t = 110$ 美元，$S_t = 100$ 美元，这意味着事前损失为 10 美元，$z_t = 1.1$。因此，当 z_t 由 1.0 增加到 1.1 时（即更大的事前损失），额外损失所遭受的痛苦现在应用 $\lambda(z_t) = 2.0 + 10(0.1) = 3$ 而不是 $\lambda = 2.0$ 来进行惩罚。更为"现实"的唯一要求是，z_t 应非常缓慢地移向 S_t，从而当 S_t 增加（或减小）时，Z_t 也会增加（或减小），但变化的幅度比 S_t 小。

$$z_{t+1} = \eta\left(z_t \frac{R}{R^*_{t+1}}\right) + (1-\eta) \qquad (1 > \eta > 0) \tag{27}$$

其中，R 为常数，η 是缓慢程度的一个度量指标。η 现在是可以变化的。$\eta = 0$ 意味着 $z_{t+1} = 1$，因此 S_t 与 Z_t 之间是一一对应关系；$\eta = 1$ 意味着 z_{t+1} 的反应很缓慢；对 $\eta \neq 1$ 而言，当 $R^*_{t+1} > R$ 时，z_{t+1} 相对于 z_t 有所下降。

最优化

跨期最优化问题为

$$\max E_t\left\{\sum_{t=0}^{\infty}\left[\theta^t\left(\frac{C_t^{1-\gamma}}{1-\gamma}\right) + b_0 \theta^{t+1} \bar{C}_t^{-\gamma} v(X_{t+1}, z_t)\right]\right\} \tag{28}$$

其中，$X_{t+1} = S_t(R^*_{t+1} - R^*_{f,t})$，"财富"约束为（无劳动收入）

$$W_{t+1} = (W_t - C_t) R^*_{f,t} + S_t(R^*_{t+1} - R^*_{f,t}) \tag{29}$$

我们来看看校准后的模型能否拟合真实数据中的典型事实，如股权溢价之谜等。一些参数使用的是历史平均值，一些参数是基于行为研究来确定取值，其他参数的选择是为了与数据中的某些性质相匹配。我们可以用数值方法来求模型的基准解，所使用的消费增长率为 $g_c = 1.84\%$ p.a，$\sigma_c = 3.79\%$，（对消费的）风险规避程度为 $\gamma = 1.0$，时间偏好率为 $\theta = 0.98$。这些参数确保了均衡无风险利率 $R_f^* - 1 = 3.86\%$（参见（31））等于数据中所发现的利率，从而"无风险

利率之谜”不存在。

k 在一个范围内取值，k 值的逐渐增加意味着，如果存在事前损失，那么随后损失所带来的痛苦会增加。例如，当 $k=3$ 时，平均损失规避程度大约为 2.25。假设 z_t 在开始的时候为 1，股票市场的价格下跌了 10%。因此，当 $\eta=1$ 时（举个例子），z_t 会向 1.1 运动，任何额外损失都得用 $2.25+3\times0.1=2.55$ 来进行惩罚，痛苦程度只有轻微的增加（参见（25））。但是，当损失所造成的痛苦为 $k=50$ 时，上述情形意味着，如果股市价格下跌 10%，那么（此时 $\eta=1$）任何额外损失都得用权重（$2.25+5=$）7.25 来进行惩罚。正如我们在下面将要看到的那样，随着 k 的增加，拟合的股权溢价也会增加，这使得拟合的股权溢价非常接近于真实数据中所观察到的股权溢价。常数 R 的设定要保证 z_t 的无条件均值等于 1。拟合结果是针对 b_0 的各种取值而言的，因为我们对这个参数的可能取值没有任何先念信息。

损失规避参数 λ 的取值为 2.25。参数 η 控制了 z_t 中的持续性，从而控制了价格红利比中的持续性：选择 $\eta=0.9$（即反应缓慢）是为了使所拟合的价格红利比存在和真实数据一样的自相关性。

数值求解程序相当复杂，还要用到迭代技巧，因为状态变量 z_{t+1} 是红利价格比和 ε_{t+1} 的函数，且函数形式为 $z_{t+1}=h(z_t,\ \varepsilon_{t+1})$。进行蒙特卡罗模拟，$\varepsilon_{t+1}$ 的 10000 次抽取就给出了序列 z_{t+1}，这反过来又可以用来模拟收益和红利价格比。

结果：“经济 I”

当红利和消费都服从相同的随机过程时，与平均股权溢价及其波动性有关的结果不是很理想。当 $b_0=2$ 且 $k=3$ 时，拟合的股权溢价为 0.88% *p. a.*（s. d. =5.17% *p. a.*），而真实数据中的结果为 6.03% *p. a.*（s. d. =20.02% *p. a.*）。即使增加到 100（给效用函数中事前损失部分赋予更大的权重）、$k=50$（即对任意事前损失而言损失规避程度更高），所拟合的平均股权溢价也只上升到 3.28% *p. a.*（*s. d.* =9.35% *p. a.*）。

结果：“经济 II”

我们现在来讨论“经济 II”，红利服从如下随机过程

$$d_{t+1}-d_t=g_d+\sigma_d\varepsilon_{t+1}$$

其中，$g_d=1.84\%$ *p. a.*，$\sigma_d=12\%$ *p. a.*，消费冲击与红利增长之间的相关系数取值为 0.15。表 2 给出了模型所产生的“感兴趣的统计量”和相应的实证结果。

当 $b_0=2$ 且 $k=3$ 时，模型产生的股权溢价为 2.62% *p. a*（s. d. =20.87%），这一结果比“经济 I”中的结果好，但仍然比实证结果 6.03% *p. a.*（s. d. =20.02%）低。如果我们增加由事前损失所带来的痛苦以至于 $b_0=2$，但 $k=10$，那么模型给出的股权溢价为 5.02% *p. a.*（*s. d.* =23.84%），这与真实数据中所观察到的结果非常接近，但损失规避的平均水平只有轻微增加，从 2.25 增加到 3.5。

股票收益的解可以写为

$$R_{t+1}=\frac{1+f(z_{t+1})}{f(z_{t+1})}e^{g_d+\sigma\varepsilon_{t+1}}$$

因此，当我们考虑 $\sigma_d=12\%$（$\sigma_c=3.79\%$）的红利过程时，股票收益的波动率要大一些。从直觉来看，更大的波动率是因为，如果存在一个正的红利创新（dividend innovation），那么这会产生更高的股票价格和更高的股票收益。但是，这也增加了事前收益，所以投资者的风险规避程度会有所降低，风险规避程度的降低会减小红利的贴现率，从而产生更高的价格和更大的收益运动（波动率）。（相反的结果适用于负的红利创新，如果存在事前损失，那么结果就会多一个“棘手的部分”。）因为平均收益有着更大的波动率且投资者预期会遭受更多的损失，所以损失规避要求一个更高的股权溢价。注意，如果没有事前结果的假设（即令 z_t 等于 0），那么当 $b_0=2$ 时，模型会产生一个非常小的股权溢价，约为 2%（s. d. =12%）；当 b_0 等于 100

时，股权溢价也只上升到2.88%（s.d. =12%）。

表 2　　经济 II 中模型所产生的统计量和相应的实证结果

（Barberis、Ming 和 Tano（2001））　　（（%）=回归的 R^2）

	模型所给出的值		实证结果
	$b_0=2$，$k=3$	$b_0=2$，$k=10$	
超额股票收益（股权溢价）			
均值	2.62	5.02	6.03
标准差	20.87	23.84	20.02
夏普比	0.13	0.21	0.30
平均损失规避	2.25	3.5	—
价格红利比			
均值	22.1	14.6	25.5
标准差	2.25	2.5	7.1
收益的自相关性			
滞后 -1	-0.07	-0.12	0.07
-2	-0.03	-0.09	-0.17
-3	-0.04	-0.06	-0.05
-4	-0.04	-0.04	-0.11
-5	-0.02	-0.03	-0.04
Corr $\{(P/D)_t, (P.\ D)_{t-k}\}$			
$k=1$	0.81	0.72	0.70
$k=2$	0.53	0.28	0.45
$k=3$	0.35	0.20	0.40
回归 $R_{t,t+k}=\alpha_k+\beta_k\ (D/S)_t$			
β_1	4.6（2%）	4.4（6%）	4.2（7%）
β_2	8.3（4%）	7.5（10%）	8.7（16%）
β_3	11.6（5%）	9.7（12%）	12.3（22%）
β_4	13.7（6%）	11.5（14%）	15.9（30%）

当我们输入红利创新与消费增长率之间的低相关系数0.15时（与真实数据中的0.1接近），模型就变得可用了。模型能产生消费与股票收益之间的低相关性。这是因为收益会对红利方面的消息和由收益变化引起的风险规避程度变化产生反应。在这个模型中，红利方面的消息和由收益变化引起的风险规避程度变化都主要受红利冲击所驱动，红利冲击与消费之间的相关性很弱（因此，模型中的收益与消费只是微弱相关，真实数据也是如此）。消费增长与股票收益之间的这种弱相关性在 Campbell 和 Cochrane（2000）的习惯持续模型中没有任何体现，在他们的模型中，风险规避程度的改变（从而收益的改变）是由消费所驱动的，这意味着这些变量之间存在很强的相关性。

风险规避程度的缓慢变化也会产生长期收益可预测性，因为投资者对事前收益或损失的反应很缓慢（参见表2）。正的红利创新导致股票价格上涨，从而使风险规避程度降低，进而使股票价格进一步上涨。现在，价格红利比已经很高，投资者的风险规避程度有所降低，投资

者的合意收益率也会降低。因此，价格红利比有助于未来收益的预测，这与 Campbell 和 Shiller（1988）的实证结果一致。因为事前收益或损失存在，所以投资者的风险规避程度也会随时间而发生变化，模型预测的收益也会随着时间而发生变化，这就导致了股票价格的可预测性。模型中的价格红利比也存在很强的持续性和很高的自相关性（表2）。

模型也会产生具有负自相关性的收益，因为高价格（收益）会导致更低的风险规避倾向，从而降低未来收益。收益的负自相关性意味着长期均值反转效应存在（如 Poterba 和 Summer（1988），Fama 和 French（1988a），Cochrane（2001），如前面章节所述）。

无风险利率又会怎样呢？在这个模型中，无风险利率是从效用函数的前景理论部分推导出来的，它由标准的欧拉方程给出：

$$1=\theta R_f^* E_t\left[(\bar{C}_{t+1}/\bar{C}_t)^{-\gamma}\right] \tag{30}$$

$\bar{C}_t$ 是总消费，在均衡状态下有

$$\ln R_f^* = -\ln\theta+\gamma g_c-(\gamma^2/2)\,\sigma_c^2 \tag{31}$$

无风险利率的均值和波动率主要由 γg_c 和消费增长的波动率来决定，其中，$g_c=1.84\%$，$\sigma_c=3.79\%$。因此，模型意味着无风险利率的平均值较小，波动率也较小。尽管 R_f 的欧拉方程是由效用函数的前景理论部分推导出来的，后者还是允许风险资产的收益有更高的波动率。欧拉方程为

$$1=\theta E_t\left[R_{t+1}^*(\bar{C}_{t+1}/\bar{C}_t)^{-\gamma}\right]+b_0\theta E_t\left[\hat{v}(R_{t+1}^*,z_t)\right] \tag{32}$$

风险资产的平均收益取决于相对于事前损失的财富变化和“标准的”总消费增长率。

在表2中要注意的是，模型能解释价格红利比的平均值，但没有解释其波动率，模型的标准差为2.25%～2.5%，而实际数据中的标准差为7.1%。其他状态变量（如消费习惯）的加入会增加模型中价格红利比的波动率，但这需要更为深入的研究。

是损失规避参数 λ 还是事前损失参数 z_t 产生了“经济 II”中的核心结果？收益的解（$z_t=0$）为

$$R_{t+1}^*=\frac{P_{t+1}+D_{t+1}}{P_t}=\frac{1+f_{t+1}}{f_t}e^{g_t+\sigma_d\varepsilon_{t+1}}$$

其中，f_t 是价格红利比，它现在不依赖于 z_t。R_{t+1} 的波动率现在只依赖于对数红利增长率的波动率 σ_d（$\sigma_d=12\%$），但这与实际数据不是很一致。因此，为了改变损失规避的程度，为了使预期收益比现金流出现更大的变动（使贴现率发生变化），在效用函数中需要事前损失的出现。因此，模型的关键因素是收益必须比现金流有着更大的变动。然而，要注意的是，损失规避的改变可能不是这一结果的唯一原因。人们对风险的认知可能有了改变，或对红利消息存在“过度反应”，或对模型中的重要参数（如红利增长率的均值和波动率）进行学习，这些问题都需要进一步研究。

前景理论模型背后的基本思想与 Campbell－Cochrane（2000）习惯持续模型非常相似（但不同）。两种模型都要求效用函数取决于一个状态变量，这个状态变量是与过去相比较的一个变量。在 Campbell－Cochrane 的模型中，状态变量是超额消费 $S_t=(C_t-X_t)/X_t$，当 C_t 接近于 X_t 时（即 C_t 接近于习惯消费），风险规避程度的变化最为敏感。在 Barberis 等人的模型中，状态变量是相对于事前收益或损失的当前股票收益（即 z_t），重要的就是这个变量。两个模型都假设变量 S_t 和 z_t 的变化很缓慢，这种缓慢变化是为了“拟合”所观察的红利价格比的持续性特征，红利价格比的持续性导致了股票价格的长期可预测性。

两个模型中的风险规避程度都随时间而变化，都依赖于各自的状态变量。Campbell－Cochrane 在效用函数中只使用了消费（相对于习惯的消费）变量，而 Barberis 等的效用函数依赖于消费和（事前）收益（即财富变化）。Barberis 等的模型具有较小的 λ 值（$=1.0$），但是为了拟合数据中的典型事实，他们的模型需要在最大值函数中增加一个（事前）损失规避项。Campbell－Cochrane 模型所要求的风险规避程度要更高一些（风险规避程度取决于相对于以往习惯的消费水平），但效用函数“更为传统”和“简洁”。从宽泛意义上讲，因为它们都是校

准后的模型，基本上都能拟合典型现实，所以你更喜欢哪个模型部分地取决于你认为哪个模型的假设更为现实一些。但这会使我们远离规范经济学王国。这两个模型都是非常有趣的模型，模型的有用扩展是引入学习行为等其他要素，甚至可以引入某种类型的非理性交易者。

读者必须注意的是，许多行为模型考虑了使用公共信息所导致的收益系统性预测误差（Barberis、Shleifer 和 Vishny（1998）），考虑了在开始时对基于私人信息预测结果的过于自信，与上述模型一样，公共信息的到达对此有一定的减缓作用（Daniel、Hirshleifer 和 Subrahmanyan（1998））。一个重要的争论是如何“统一”这些模型，因为模型的假设条件通常是为了解释特殊现象（如价格反转之后的不足反应）而量身定做的，模型可能无法解释其他现象（如共变）。另一方面，为了解释一些特殊的“异像”（如股权溢价），人们对标准的跨期预期效用最大化模型进行了“修正”。模型检验中所使用的不同技巧（如对参数的标准假设检验方法与校准和模拟方法）也使得争论没有定论。Barberis 和 Thaler（2003）对这些问题进行了很好的综述。

19.9 小结

* 近年来，行为金融领域的文献出现了爆炸性增长，试图解释观察到的异像（如封闭式基金折价、动量利润）和“典型事实”（如股权溢价、股票收益的可预测性）的各种模型由此而生。模型通常不能用回归方法和规范的假设检验来进行检验，而应使用某种形式的校准和模拟方法。

* 一些行为模型对标准效用函数进行了修正，在效用函数中加入了消费之外的其他变量，如由股票市场波动所引起的财富变化，模型中的投资者可能遭受“损失规避”。

* 一些行为模型集中关注理性交易者或理性套利者与噪声交易者之间的交互作用，模型（经常）假设噪声交易者的投资决策是基于股票价格的历史运动。“信息观察者”和动量交易者之间与未来现金流有关的私人信息和公共信息的到达时间对短期动量利润和长期价格反转及可预测性至关重要。“风格投资”模型可以很好地解释股票收益的横截面相关性（股票的现金流之间不相关）。

* 在 DeLong 等的模型中，理性代理人和噪声交易者都存在，他们都追求期末财富的最大化，但噪声交易者对未来收益可能持非理性的乐观或悲观态度。噪声交易者风险意味着，均衡价格可能持久地低于内在价值（即使噪声交易者与理性交易者有着相同的预期）。

* Shleifer - Vishny 模型强调资金借入成本（为了购买股票）与公司未来前景方面的适时信息之间的交互作用。错误定价越严重，公司投资决策就要花费更长的时间才能在市场上取得成功。这会促使经理人在选择投资项目时采取短期主义策略。

* 一些噪声交易者模型，如 Kirman（1993a），依赖于观点的传染和转换来产生快速的大幅价格运动，这基本上与观测数据一致。

* 前景理论假设个体关心收益、损失和消费水平，投资者倾向于损失规避。在 Barberis、Ming 和 Tano（2001）的模型中，个体追求终生效用的最大化，效用函数是消费和财富变化的函数，校准后的模型（使用消费和红利的观测数据）可以用来进行模拟。模型的预测结果能“拟合”许多典型事实，包括无风险利率的低水平值和低波动性、股票收益的长期可预测性和相对较高的股权溢价（但比实际数据中所观察到的股权溢价低）。

* “行为模型”领域已有很大的进展，这会持续影响我们对一些现象潜在原因的看法，如过度波动性、股权风险溢价、股票收益的可预测性和金融异像等。

附录 I　DeLong 等的噪声交易者模型

DeLong、Shleifer、Summers 和 Waldmann（1990b）的基本模型是一个两期世代交叠模型。在第一期中不存在消费和劳动供给决策，代理人的投资来源是外生的，他们的唯一决策就是在第一期（即在他们年轻的时候）选择投资组合来使期末财富的期望效用达到最大。“年老者”把风险资产卖给“新的年轻”一代，使用安全资产收益来购买消费品。安全资产“s”具有完全供给弹性。风险资产的供给是固定的，标准化后的风险资产数量为 1。两种资产都要支付一个已知的实际红利 r（$=$无风险利率），因此基本面风险不存在。1 单位安全资产可以购买 1 单位消费品，从而安全资产的实际价格为 1。

市场中噪声交易者（NT）的比例为 μ，理性套利者（SM）的比例为（$1-\mu$）。SM 在 $t+1$ 时刻清楚地知道风险资产收益的分布。NT 可能是“乐观主义者”或“悲观主义者”，对真实价格的分布存在认知偏差。NTV 对预期价格的平均错误认知用 ρ^* 表示，在任何时点，实际错误认知为

$$\rho_t \sim N\ (\rho^*,\ \sigma^2) \tag{A1}$$

每个代理人都追求效用最大化，效用函数是期末财富的函数，绝对风险规避系数为常数，

$$U = -\exp\ (-2\gamma W) \tag{A2}$$

如果风险资产的收益服从正态分布，那么最大化（A2）等价于最大化

$$EW - \gamma\sigma_w^2 \tag{A3}$$

其中，EW 是期末财富的均值，γ 是绝对风险规避系数。因此，SM 通过最大化如下目标函数来选择风险资产的持有数量 λ_t^s

$$E\ (U)\ = c_0 + \lambda_t^s\ [r + {}_tP_{t+1}^e - P_t\ (1+r)]\ - \gamma\ (\lambda_t^s)^2\sigma_{p,t+1}^2 \tag{A4}$$

其中，c_0 为常数，${}_t\sigma_{p,t+1}^2$ 是价格的提前一期（one period ahead）条件期望方差

$${}_t\sigma_{p,t+1}^2 = E_t\ [P_{t+1} - E_tP_{t+1}]^2 \tag{A5}$$

NT 的目标函数与 SM 相同，但他们的预期收益要多 $\lambda_t^n\rho_t$ 项（当然，在（A4）中要用 λ_t^n 来代替 λ_t^s）。这类目标函数与只有两种资产（其中一种资产是安全资产）的均值方差模型有着相同的形式。在（A4）中令 $\partial E\ (U)\ /\partial\lambda_t = 0$ 就得到 SM 和 NT 风险资产的均值方差需求函数

$$\lambda_t^s = E_tR_{t+1}/2\gamma\ ({}_t\sigma_{p,t+1}^2) \tag{A6}$$

$$\lambda_t^n = \frac{E_tR_{t+1}}{2\gamma\ ({}_t\sigma_{p,t+1}^2)} + \frac{\rho_t}{2\gamma\ ({}_t\sigma_{p,t+1}^2)} \tag{A7}$$

其中，$E_tR_{t+1} = r + {}_tP_{t+1}^e -\ (1+r)\ P_t$。NT 的需求部分地取决于他们对预期收益的反常认识（在 ρ_t 中所有反映）。因为“年长者”将风险资产卖给年轻人，且风险资产的固定供给为 1，所以我们有

$$(1-\mu)\ \lambda_t^s + \mu\lambda_t^n = 1 \tag{A8}$$

因此，使用（A6）和（A7）就得到均衡定价方程

$$P_t = \frac{1}{1+r}\ (r + {}_tP_{t+1}^e - 2\gamma_t\sigma_{p,t+1}^2 + \mu\rho_t) \tag{A9}$$

模型均衡是一种稳定状态，P_{t+1} 和 P_t 的无条件分布相同。递归求解（A9）就 s 得到

$$P_t = 1 + \frac{\mu\ (\rho_t - \rho^*)}{1+r} + \frac{\mu\rho^*}{r} - \frac{2\gamma_t\sigma_{p,t+1}^2}{r} \tag{A10}$$

在（A10）中只有 ρ_t 是变量，所以

$${}_t\sigma_{p,t+1}^2 = \sigma_{p,t+1}^2 = \mu^2\sigma^2/\ (1+r)^2 \tag{A11}$$

由（A1）有 $\rho_t - \rho^* = N\ (0,\ \sigma^2)$。将（A11）代入（A10）就得到正文中的价格方程

$$P_t = 1 + \frac{\mu\ (\rho_t - \rho^*)}{1+r} + \frac{\mu\rho^*}{r} - \frac{2\gamma\mu^2\sigma^2}{r\ (1+r)^2} \tag{A12}$$

附录 II　Shleifer - Vishny 的短期主义模型

我们在这个附录里正式地给出 Shleifer 和 Vishny（1990）的短期主义模型。在他们的模型中，长期资产比短期资产有着更大的定价偏差。正如正文所解释的那样，这会促使公司管理人员选择那些短期内能产生现金流的投资项目，他们这样做的目的是为了规避严重的错误定价和被接管的风险。

三个时期分别为 0、1 和 2，公司要么选择“短期”投资项目（项目在第 2 期的偿付为 V_s 美元），要么选择“长期”投资项目（项目在第 2 期的偿付为 V_g 美元）。两个项目的最大不同在于，短期投资项目的真实价值在第 1 期就可以知道，而长期投资项目的真实价值必须等到第 2 期才会明朗。因此，套利交易者关心的不是投资项目产生现金流的时机，而是错误定价的时机，尤其是错误定价在市场中得到反映的时点。无风险利率为 0，所有投资者都是风险中性投资者。

假设有两类交易者，噪声交易者（NT）和理性套利者（SM）。NT 在 $t=0$ 时刻对两类项目的偿付 V_i（$i=s$ 或 g）持悲观（$S_i>0$）或乐观的态度。因此，两类项目都面临系统性的乐观情绪或悲观情绪。我们只讨论悲观情绪情形（即 NT 持悲观情绪）。NT 对项目 i（$i=s$ 或 g）的股本需求为：

$$q\ (\mathrm{NT},\ i)\ =\ (V_i - S_i)\ /P_i \tag{A1}$$

在牛市行情中，q 应等于（$V_i - S_i$）$/P_i$。SM 面临 b 美元的借贷约束，毛利率 $R>1$（即大于 1 + 无风险利率）。因为 SM 属于风险中性，所以他们不关心 b 美元到底该投资到哪类项目上，他们的需求曲线为：

$$q\ (\mathrm{SM},\ i)\ = n_i b/P_i \tag{A2}$$

其中，n_i 是 SM 在资产 i 上的投资数量。每种资产（i）的供给为 1 个单位，因此均衡时候有

$$1 = q\ (\mathrm{SM},\ i)\ + q\ (NT,\ i) \tag{A3}$$

使用（A1）和（A2）就可以得到资产的均衡价格

$$P_i^e = V_i - S_i + n_i b \tag{A4}$$

假设 $n_i b < S_i$，从而两种资产在 $t=0$ 时刻都存在定价偏差。如果 SM 在 $t=0$ 时刻投资 b 美元，那么他能获得 b/P_s^e 份短期资产。每份短期资产的偿付 V_s 在 $t=1$ 时刻能得以实现。第 1 期的总偿付为 V_s（b/P_s^e）。第 1 期的借贷成本为 bR，净收益 NR_s 为

$$NR_s = \frac{V_s b}{P_s^e} - bR = \frac{bV_s}{V_s - S_s + n_s b} - bR \tag{A5}$$

我们在这里用到了方程（A4）。SM 在 $t=0$ 时刻进行长期资产投资，购买 b/P_g^e 份资产。他在 $t=1$ 期什么也不做。在 $t=2$ 期，每份资产的真实价值 V_g 得到体现，V_g 在 $t+1$ 期按照 R 进行贴现，这意味着偿付为（bV_g）/（$P_g R$）。属于 $t=2$ 期部分的数量为 bR^2，在期按照 R 进行贴现后的值为 R。因此，第 1 期的净收益 NR_g 为

$$NR_g = \frac{bV_g}{P_g R} - bR = \frac{bV_g}{R\ (V_g - S_g + n_g b)} - bR \tag{A6}$$

（A5）和（A6）之间的唯一差异是，在（A6）中，持有（存在定价误差）长期资产的收益要贴现到 $t=1$ 期，因为它的真实价值直到 $t=2$ 期才能得到体现。在均衡状态下，长期资产和短期资产为套利者所带来的收益必须相等（$NR_g = NR_s$），因此，由（A5）和（A6）有

$$\frac{V_g/R}{P_s^e} = \frac{V_s}{P_s^e} \tag{A7}$$

因为$R<1$，所以在均衡状态下，长期资产比短期资产有着更为严重的价格低估（当NT持悲观情绪时，$S_i>0$）。错误定价之所以产生是因为短期资产的偿付在第1期就已经明确，而长期资产的偿付要在第2期才能明确。对短期资产而言，价格在第1期就向内在价值V_i方向运动，而长期资产的价格要到第2期才能向内在价值方向运动。因此，长期资产的内在价值V_g必须贴现到第1期，这种“借入成本”降低了长期资产的持有收益。

第20章

期限结构理论

目的

* 分析零息债券、付息债券、即期收益率、连续复利即期收益率、持有期收益率和到期收益率。

* 说明理性定价公式如何用于债券价格的确定。

* 考察各种期限结构模型，包括（纯）预期假说、流动偏好假说、市场分割假说和偏好习惯理论。

* 解释收益曲线的形状。

人们对长期利率和短期利率之间关系感兴趣的一个原因是，绝大多数国家的中央银行有时会（或试图）改变短期利率，他们将短期利率视为作用于实体经济的一种工具，通过短期利率来影响通货膨胀率。货币当局通过公开市场操作（如买卖债券）或威胁这样做来实现这一目标。短期利率的改变（但通货膨胀预期没有改变）可能会对实际存货持有量和消费者的消费（尤其是耐用品的消费）产生影响。短期利率可能影响期限较长的政府债券的利率：这就是收益曲线或期限结构关系。如果政府贷款利率对公司债券利率产生影响，那么公司债券利率就可能影响公司在厂房和机器设备上的实际投资金额。因此，货币政策通过直接或间接的方式来对实际经济活动产生影响（Cuthbertson 和 Nitzsche（2001a））。

如果中央银行的利率政策旨在抑制通货膨胀，那么短期利率的上升可能导致对未来通货膨胀预期的降低，也可能导致未来短期利率的降低，因此短期利率的上升可能使长期利率降低——这种情形下的收益曲线是一种向下的收益曲线。另一方面，如果名义短期利率随着通货膨胀率的增加而增加，那么长期利率可能是递增的形状，因为人们预期将来的通货膨胀率会更高，因此这种情形下的收益曲线是一种向上的收益曲线。国内短期利率预料不到的变化可能对资本流动和汇率产生影响（这种联系根源于更大的不确定性），但是，如果可以操作的话，利率政策可能对价格的竞争程度、贸易净额（出口减进口）、就业和产出水平等产生影响。我们将在第 24 章讨论短期利率和汇率的这种联系。

金融经济学家对债券价格和利率行为很感兴趣，因为他们可以用债券价格和利率来对与市场参与者和市场效率有关的各种行为假设进行检验。债券价格可以用来检验 EMH 的有效性，由此得到的结果可以用来与以股票收益和股票价格为基础的检验结果（像前面相关章节讨论的那样）进行比较。EMH 下的名义股票价格等于预期红利支付的 DPV 值。可以类似地将政府债券的名义价格视为名义息票支付的 DPV 值。然而，众所周知，名义息票支付是确定的，理性预期假设下债券价格变化的唯一来源是与远期单期利率有关的消息（如折现因子）。在债券市场上，对 EMH 的检验倾向于以“收益率”的行为为中心，而不是以价格为中心。比如，Shiller 的方差边界法检验（Variance Bound Test）将 P_t 和完全预见（perfect forsight）股票价格

P_t^* 进行比较，而 Var 方法则将 P_t 和理论的 P_t' 进行比较。类似地，在债券市场中，我们可以将完全预见收益率 R_t^* 或完全预见收益率差（yield spread）与真实值 R_t 和 S_t 分别进行比较。我们将在下一章中讨论这些问题，在此之前，我们在这一章中交代一些基本概念并发展期限结构理论。

20.1　价格、收益与 RVF

债券所提供的投资机会可以用持有期收益率、即期收益率和到期收益率来表示，因此债券的“收益”可以用多种方法来进行定义。在这一节中，我们首先阐述这些指标之间的关系，然后在预期“收益”的各种假设下，在 EMH 的基础上来阐述债券市场参与者的各种行为假设。

“大众型”（plain vanilla）付息债券在债券购买时就已经确切地知道未来各期的收入流：（名义的）息票支付为 C_{t+i}。在绝大多数情形中，任何时期的 C_{t+i} 都是常数，保留下标有时有助于问题的说明。绝大多数债券，与股票不一样，在将来的某个给定日期（$=t+n$）能以众所周知的价格 M_n 赎回，这个价格也称为票面价值（par value）、赎回价格或到期值。有些债券，尽管支付息票，但永远不能赎回，这种债券就是众所周知的永续年金（如 $2\frac{1}{2}\%$ 公债，英国政府发行）。

短期债券（bill）（如国库券）没有任何息票支付，但赎回价格在债券发行日是固定的。因此，短期债券的收益就是发行价格（或购买时的市场价格）与赎回价格之间的差（用百分比表示）。短期债券一般采用贴现的方式进行发行（也就是说，发行价格低于赎回价格），目的在于保证短期债券持有者在债券的有效期内能获得正的收益。因此，短期债券也被称为纯贴现债券或零息债券。市场中交易的绝大多数短期债券的期限都很短（也就是说，这些债券在发行后三个月、六个月或一年到期）。另一方面，付息债券的期限一般超过 1 年，在非常活跃的金融市场中付息债券的期限为 5 ~ 15 年。在这一章里，我们只考虑（不可赎回）“固定息票支付”的政府债券和短期债券，并假设这些债券没有任何违约风险（也就是说，我们不考虑公司债券）。

因为付息债券与股票在很多方面是相似的，因此我们可以将股票市场中所导出的许多分析思想、分析理论和计算公式应用于对债券市场的分析。

即期收益率/即期利率

即期收益率（或即期利率）指在给定时期中以给定的无风险利率进行借贷资金所获得（或支付）的收益率。比如，假设你以 $r^{(1)}$ 的利率贷出资金（比方说，贷给银行），从 $t=0$ 到 $t=1$ 的一年期贷款的利率就为 $r^{(1)}$。对 0 时刻的 P 美元投资而言，一年后银行将支付 $M_1=P(1+r^{(1)})$ 美元。假设银行“两年期货币”用年复利表示的利率为 $r^{(1)}$，那么 P 美元投资在两年后将增值到 $M_2=P(1+r^{(2)})^2$ 美元。因此，今天挂牌的 n 年期即期利率假定今天投资的 P 美元在未来 n 年中的收益率是“锁定的”。

根据原理，即期利率序列可以根据市场上所能观察到的纯贴现债券价格来进行计算。如果期限为 $n=1, 2, \cdots$ 的债券在市场上所能观察到的价格分别为 $P_t^{(1)}$，$P_t^{(2)}$，…，那么可由 $(1+r_t^{(n)})^n=M/P_t^{(n)}$ 来导出每一期的即期利率。

看待即期利率的一种等价方法是，即期利率可以用作贴现利率来计算无违约风险情形下未来某个特定日期的货币增值额。如果你在 n 年后将获得 $M=1$ 美元，那么这笔资金的 DPV 值为 $M/(1+r^{(n)})^n$ 美元（事实上，这就是 n 年期零息债券在今天的“公平”价格），

$$p_t^{(n)}=\ln P_t^{(n)}=\ln M-n\ln(1+r_t^{(n)}) \tag{1}$$

我们也可以用连续复利利率（参见第一章）来表示价格，连续复利利率定义为 $rc_t = \ln(1+r_t)$，因此

$$p_t^{(n)} = \ln P_t^{(n)} = \ln M - n\,[rc_t] \quad \text{或} \quad P_t^{(n)} = M\exp(-rc_t n) \tag{2}$$

实践中，在较长期限范围（如超过一年）的最后时刻，（贴现）短期债券或纯贴现债券通常不存在。然而，较长期限的即期利率可以用付息债券的数据来进行逼近（尽管我们在这里并不需要考虑这些细节，参见 McCulloch（1971，1990））。

如果我们拥有 n 期付息债券，由市场决定的所有期限的即期利率都存在，那么债券的市场价格为

$$P_t^{(n)} = \frac{C_{t+1}}{(1+r_t^{(1)})} + \frac{C_{t+2}}{(1+r_t^{(2)})^2} + \cdots + \frac{C_{t+n}+M_{t+n}}{(1+r_t^{(n)})^n} = \sum_{i=1}^{n} V_i \tag{3}$$

其中，M = 到期（赎回）值，$V_i = \frac{C_{t+i}}{(1+r_t^{(i)})^i}$，$i = 1, 2, \cdots, n-1$，$V_n = \frac{C_{t+n}+M_{t+n}}{(1+r_t^{(n)})^n}$。债券的市场价格等于未来息票（和到期值）的 DPV 值，贴现率为即期利率。如果上述公式不成立，那么进行息票剥离（coupon stripping）就能获得无风险套利利润。为了说明这个问题，考虑两期付息债券，假设该债券的市场价格 $P_t^{(2)}$ 小于 $V_1 + V_2$。支付为 C_{t+1} 和 $C_{t+2} + M_{t+2}$ 的付息债券在当前时刻的市场价格分别为 V_1 和 V_2。如果 $P_t^{(2)} < V_1 + V_2$，那么人们就会购买两年期付息债券，同时卖出“息票支付”的权利给其他市场参与者，承诺在第一年和第二年末分别支付 C_{t+1} 和 $C_{t+2} + M_{t+2}$。如果零息债券的定价是正确的，那么这些权利在今天的售价分别为 V_1 和 V_2。因此，在瞬间就可以获得无风险利润（$V_1 + V_2 - P_t^{(2)}$）。在有效金融市场上，两年期付息债券需求的增加使得 $P_t^{(2)}$ 升高，与此同时，息票的售卖将降低一年期和两年期零息债券的价格。因此，这种无风险套利使得（3）式重新成立。最后要注意的是，债券的“扁平收益”（flat yield）为 C/P_t，债券的息票率为 C/M。

持有期收益率，HPR

如果在 t 到 $t+1$ 之间持有付息债券，那么总收益包括资本利得和息票支付两个部分。就债券而言，“收益”的测度指标是众所周知的（单期）持有期收益（或持有期收益率）：

$$H_{t+1}^{(n)} = \frac{P_{t+1}^{(n-1)} - P_t^{(n)} + C_t}{P_t^{(n)}} \tag{4}$$

对零息债券而言，$C_t = 0$，HPR 就是资本利得。在上述公式中需要注意的是，n 期付息债券是一个时期之后的（$n-1$）期付息债券（在对长期债券（比如 $n > 10$ 年）的实证研究中，研究人员经常用 $P_{t+1}^{(n)}$ 来代替 $P_{t+1}^{(n-1)}$，因为对所收集的周数据、月度数据或季度数据而言，两者近似相等）。

（4）中 H_{t+1} 定义为事后的（或实际的）HPR，因为 t 时刻投资者并不知道 P_{t+1}，所以投资者会对 P_{t+1} 形成预期，因而也会对 HPR 形成预期（注意，在时刻通常都能知道未来各期的息票支付额）。

到期收益率（赎回收益率）

对付息债券而言，市场上挂牌的利率是到期收益率（YTM）。投资者知道债券当前的市场价格 P_t、息票支付现金流 C（每年支付一次）、债券的赎回价值（或到期值）（$=M$）和到期日 n。现在假设不同时期的息票支付在 t 时刻都采用相同的贴现率 $1/(1+y_t)$。注意 y_t，也就是到期收益率（YTM），有一个下标 t，因为它会随着时间的演变而发生改变（但在任何一个时期的 DPV 公式中，它是不会发生改变的）。如果令息票支付的 DPV 值与债券当前的市场价格相等，那么我们有

$$P_t = \frac{C}{(1+y_t)} + \frac{C}{(1+y_t)^2} + \cdots + \frac{C+M}{(1+y_t)^n} \tag{5}$$

债券可以视为一种投资，今天的投资总额为 P_t，在将来的各个时期获得一个已知的现金流（C 和 M）。使（5）中左右两边相等的常数 y_t 就是这一投资的“内部收益率”。对债券而言，y_t 称为（n 期）债券的到期收益率或赎回收益率。显然，在市场价格发生变化的任何时刻我们都必须计算 y_t。财经媒体一般会报道债券价格、息票率和到期收益率。值得注意的是，由市场变量（P_t，C，M，n）可以导出到期收益率，在经济意义下，到期收益率无法决定市场价格。事实上，在（3）中，可以用即期利率 r_{t+i} 来确定市场价格 P_t，金融界又根据上述 P_t 值通过（5）式来计算 y_t。市场挂牌 YTM 的原因在于，对给定期限的每一种债券而言，这种方式只需提供一个数据（而（3）中涉及的是即期利率序列）。一般而言，n 期债券的到期收益率和 m 期债券的到期收益率在任何时点都不会相同，因为每一种债券的息票支付 C 不同，息票支付在不同时段（也就是 n 和 m）上的贴现值也不同。从（5）容易看出，债券价格和到期收益率的运动方向是相反的，对到期收益率 y_t 所给定的任何变化而言，长期债券价格变化的百分比要大于短期债券价格变化的百分比。另外，到期收益率计算公式（5）可以简化为永续债券的计算公式 $P_t = C/y$（也就是说，$n \to \infty$ 时的情形）。

尽管财经媒体对到期收益率进行过广泛报道，但到期收益率有时是债券“收益”一个模棱两可的测度指标。在计算到期收益率时，人们隐含假设投资者在债券有效期内未来各期都能以常数利率 y_t 对息票支付进行再投资。为了说明这个问题，考虑由（5）给出的两期债券的到期收益率，进行简单变换之后，我们有

$$(1+y_t)^2 P_t = C\ (1+y_t)\ +\ (C+M) \tag{6}$$

左边是以常数年利率 y_t 投资 P_t 美元的最终价值（投资两年时间），右边包括两个部分：$t+2$ 时刻的支付（$C+M$）；第一年年末的息票支付 C 以利率 y_t 进行再投资后在 $t+2$ 时刻的价值。因为（5）和（6）是等价的，所以 DPV 公式假定第一笔息票支付在第一年后的再投资利率为 y_t。然而，没有任何证据能证明投资者能以常数利率 y_t 对未来各期的息票支付进行再投资。注意，这里的问题不是投资者必须对息票支付的未来再投资利率形成某种观念，而是（比如说）他们所选择的适用于20年债券的、在第9年和第10年间的再投资利率必须等于20年债券当前的到期收益率 y_t。

我们可以用另一种方式来考察用到期收益率作为付息债券收益测度指标所带来的问题。考虑两种债券，这两种债券有着不同的息票支付流 $C_{t+i}^{(1)}$ 和 $C_{t+i}^{(2)}$，但两种债券的价格、到期日和到期值都相同。根据（5），这意味着两种债券的到期收益率 y_{1t} 和 y_{2t} 不同。如果投资者的投资者组合中包含这两种债券，并且投资者相信公式（5）的正确性，那么他必须隐含地假定，在 $t+j$ 到 $t+j+1$ 期间，他能对第一种债券的息票支付以 y_{1t} 的利率进行再投资，而对第二种债券的息票支付以 y_{2t} 的利率进行再投资。但在现实生活中，$t+j$ 到 $t+j+1$ 期间的再投资利率对两种债券而言是相同的，再投资利率等于所有年份中都适用的单期即期利率（投资者在时刻并不知道该利率）。

除上述问题之外，如果持有债券直到期限结束，通常认为 YTM 是债券年平均百分比收益的一个非常有用的近似指标。然而，正因为到期收益率概念存在上述缺陷，所以基于这一指标的收益率曲线通常很难用十分清晰的方式进行解释。

20.2　期限结构理论

我们先根据单期 HPR，然后利用即期收益率来考察期限结构理论。在下一章中，我们用更为正式的方式来说明 HPR 方程是如何与长期和短期债券的即期利率产生联系的。

预期假说

如果所有投资者都是风险中性的，所有投资者都只考虑预期收益，那么无论债券期限的长

短，所有债券预期的单期 HPR（也就是说，时间长度为一个月或一个季度的 HPR）都应等于单期资产（如 1 月期国库券）的已知（无风险）收益率 r_t。

$$E_t H_{t+1}^{(n)} = r_t \quad （对所有的 n） \tag{7}$$

这就是纯预期假说（pure expectations hypothesis）（PEH），所有债券（不管期限的长短）的期限溢价（term premium）或风险溢价（risk premium）都为零。所有边际投资者都是赌徒（plungers）。例如，假设三年后到期的债券与两年后到期的债券相比，存在一个预料不到的 HPR。投资者将卖出两年期债券买进三年期债券，因此三年期债券的当前价格将上升，债券的单期 HPR 减小。两年期债券价格的变动与三年期债券价格的变动方向相反，因此，所有的持有期收益率都将相等。如果我们增加理性预期（RE）假设，那么

$$H_{t+1}^{(n)} - r_t = \eta_{t+1}^{(n)} \quad （对所有的 n） \tag{8}$$

因此，事后的超额持有期收益率序列均值为零，且与 t 时刻的信息集合（Ω_t）独立。因此，残差序列 $\eta_{t+1}^{(n)}$ 是序列不相关的（但可能出现异方差）。这就为 PEH（+RE）的检验提供了一种方法。

上述结论似乎合情合理，因为持有长期债券的收益（就某一个时期而言）是不确定的（因为 t 时刻不知道下一期的期末价格是多少），超额 *HPR* 应该依赖于某种形式的风险报酬或期限溢价 $T_t^{(n)}$。

$$E_t H_{t+1}^{(n)} = r_t + T_t^{(n)} \tag{9}$$

如果模型中没有期限溢价这一项，那么方程（9）仅仅是一个同意反复。关于期限溢价的最简单（非平凡）假设是：（i）期限溢价在整个时期内为常数；（ii）期限溢价与债券的期限无关（也就是说，$T_t^{(n)} = T$）。这种情形对应的就是预期假说（expectation hypothesis）（EH）（表 1）。显然，EH 与 PEH 有着相似的预测结果，也就是说，超额 HPR 序列无关，超额 HPR 序列与 Ω_t 相互独立。注意，此时的超额 HPR 对所有期限而言都是一个常数。对 EH + RE 和时间不变期限溢价（time invariant term premium）而言，我们可以得到如下方差不等式

$$\mathrm{Var}\ (H_{t+1}^{(n)}) \geqslant \mathrm{Var}\ (r_t) \tag{10}$$

因此，n 期债券单期 HPR 的方差要大于等于单期无风险利率的方差。与此结论等价的一个结论是，超额 HPR 的方差对不同期限的债券而言都相同。

流动偏好假设，LPH

在这一节中，我们假设期限溢价不随时间的演变而发生改变，但期限溢价依赖于债券的期限（也就是说，$T_t^{(n)} = T^{(n)}$）。例如，与期限较短的债券相比，可以认为期限更长的债券包含更多的“风险”，即使我们所考虑的是两种债券的某个固定持有期，这种看法也是合理的。这种情形很有可能发生，因为就任何给定的收益率变化而言，期限越长的债券价格变化越大。考虑如下情形：20 年期债券的一月 HPR 比 10 年期债券的一月 HPR 波动程度高。在这种情形中，为了让投资者在持有 10 年期债券的同时愿意持有 20 年期债券，投资者必然要求在 20 年期债券上获得一个更高的收益率。如果价格波动率的这种差异仅仅依赖于 n 期限之间的差异，那么这将导致一种仅仅依赖于 n 的流动性溢价（liquidity premium）。

流动偏好假说认为预期的超额 HPR 对任意给定期限而言是一个常数，但对那些期限更长的债券而言，期限溢价也将更高。也就是说，10 年期债券的预期超额 HPR 要比 5 年期债券的预期超额 HPR 高，但两者之间的差异不会随着时间的演变而发生改变。因此，（例如）与 5 年期债券相比，10 年期债券存在 1% 的预期超额收益率，在所有时期中这一差异都为 1%。当然，根据数据研究的结果表明，实际的 10 年期债券的超额收益率围绕预期 HPR 随机变动，因为在每一个时期中都存在预测误差（均值为零）。在流动偏好假说中，我们有

$$E_t H_{t+1}^{(n)} - r_t = T^{(n)} \quad (T^{(n)} > T^{(n-1)} > \cdots) \tag{11}$$

在理性预期（RE）假设下，流动偏好假说再次给出如下预测结果：HPR 序列无关且与 t 时刻的信息集合无关。因此，除了非常平凡的常数项外，运用回归分析，PEH、EH 和流动偏

好假说（LPH）的检验含义是相同的。对 PEH 而言，T 为零常数；对 EH 而言，T 为常数；对 LPH 而言，期限不同的债券有着不同的常数。因为 PEH、EH 和 LPH 彼此类似，所以我们经常将三者简单地称为 EH。我们将在下一节中见到 EH（零期限溢价或常数期限溢价）与时变期限溢价模型之间的关键性差异。

表 1　　期限结构：小结

所有收益率都是连续复利收益率。

1. 纯预期假说：PEH

（a）预期超额收益率为零

（b）对所有期限而言，期限溢价为零

$$E_t H_{t+1}^{(n)} - r_t = 0 \quad R_t^{(n)} - E_t\ (r_{t+j}'s)\ = 0$$

2. 预期假说（EH）或常数期限溢价

（a）对所有期限而言，预期超额收益率都等于一个相同的常数

（b）对所有期限而言，期限溢价都等于一个相同的常数

$$E_t H_{t+1}^{(n)} - r_t = T \quad R_t^{(n)} - E_t\ (r_{t+j}'s)\ = \Phi^{(n)}$$

其中 T 和 $\Phi^{(n)}$ 分别是单期期限溢价和期期限溢价。

3. 流动偏好假说：LPH

（a）期债券的超额收益率是一个常数，且到期时间越长，该常数值越大

（b）期限溢价是期限 n 的递增函数，

$$E_t H_{t+1}^{(n)} - r_t = T^{(n)} \quad R_t^{(n)} - E_t\ (r_{t+j}'s)\ = \Phi^{(n)}$$

其中，$T^{(n)} > T^{(n-1)} > \cdots$。

4. 时变风险（Time ~ Varying Risk）

（a）期债券的超额收益率随着 n 和时间的变化而变化

（b）期限溢价依赖于 n，且随着时间变化而变化

$$E_t H_{t+1}^{(n)} - r_t = T\ (n,\ z_t) \quad R_t^{(n)} - E_t\ (r_{t+j}'s)\ = \Phi\ (n,\ z_t)$$

其中，T（.）是关于 n 和一组变量 z_t 的函数。

5. 市场分割假说

（a）超额收益至少部分地受不同期限资产持有比例的影响

（b）期限溢价部分地依赖于不同期限资产的持有比例

$$E_t H_{t+1}^{(n)} - r_t = T\ (z_t^{(n)}) \quad R_t^{(n)} - E_t\ (r_{t+j}'s)\ = \Phi\ (z_t^{(n)})$$

其中，$z_t^{(n)}$ 是期限为 n 的资产的持有量占总资产的比例的一个测度指标。

6. 习惯偏好理论

到期时间比较接近的债券是较好的替代品，因此，这些到期时间比较接近的债券有着相似的期限溢价。

市场分割假说

市场分割假说可以视为一组资产需求方程的市场均衡解或简化形式。为了简化起见，假设我们仅拥有两种风险资产：债券 B_x 和 B_y 债券，两种资产的需求函数分别为

$$(B_x/W)^d = f_1\ (E_t H_{x,t+1} - r_t,\ E_t H_{y,t+1} - r_t,\ Z) \tag{12a}$$

$$(B_y/W)^d = f_2\ (E_t H_{x,t+1} - r_t,\ E_t H_{y,t+1} - r_t,\ Z) \tag{12b}$$

其中，W = 金融财富，Z = 影响需求的其他变量，r = 国库券的无风险收益，$E_t H_{x,t+1}$ = 期限为 x

的债券的预期持有期收益率。国库券的需求函数由预算约束的剩余部分给出

$$(TB/W) = 1 - (B_x + B_y)/W \qquad (13)$$

我们不需要考虑国库券的需求。如果我们现在假设 B_x 和 B_y 的供给是外生的，那么由（13）可以解出均衡市场利率

$$E_tH_{x,t+1} - r_t = g_1[B_x/W,\ B_y/W,\ Z] \qquad (14a)$$

$$E_tH_{y,t+1} - r_t = g_2[B_x/W,\ B_y/W,\ Z] \qquad (14b)$$

因此，不同期限的两种债券的超额 HPR 依赖于每种资产在总财富中的比例。这就是市场分割假说的基础。一般而言，（12）的需求函数包含许多自变量 Z，如收益率的方差、实际财富、通货膨胀率以及这些变量的滞后值。

习惯偏好假说

从实际效果来看，习惯偏好理论并不知道期限溢价取决于哪些因素。这种理论给我们的建议是，我们只需对到期时间相似的政府债券的收益率进行比较，然后就可以做出如下预期：超额持有期收益率应该比较接近。

20.3 预期假说

预期假说也可以用即期收益率来导出，$R_t^{(n)}$ 和 r_t 是每期利率。考虑在 n 年期（零息）债券上的 A 美元投资。投资的最终价值为

$$TV = A(1 + R_t^{(n)})^n \text{ 美元} \qquad (15)$$

这里的 $R_t^{(n)}$ 是 n 期长期债券的年（复利）利率。接下来，考虑 A 美元投资的替代性策略，利息收益来自于 n 年中单期投资的滚动收益。如果忽略交易成本，那么单期投资序列的预期最终价值 $E_t(TV)$ 为

$$E_t(TV) = A(1 + r_t)(1 + E_tr_{t+1})(1 + E_tr_{t+2})\cdots(1 + E_tr_{t+n-1}) \text{ 美元} \qquad (16)$$

这里的 r_{t+i} 是从 $t+i$ 到 $t+i+1$ 的那个时期上的利率。长期债券投资的最终价值是已知的，因为购买这种债券的人会持有该债券一直到期。一系列一年期投资的最终价值是不确定的，因为投资者必须对将来的单期即期收益率 r_{t+j} 进行猜测。然而，在风险中性条件下，这两种投资策略的最终价值相等：

$$(1 + R_t^{(n)})^n = (1 + r_t)(1 + E_tr_{t+1})(1 + E_tr_{t+2})\cdots(1 + E_tr_{t+n-1}) \qquad (17)$$

等式成立的原因在于，如果在长期债券上进行投资的最终价值大于一年期投资序列的预期最终价值，那么投资者在 t 时刻就会买入长期债券，卖出短期债券。这样操作的结果是，在最终价值给定的条件下，长期债券的当前市场价格上升，长期（即期）收益率 $R_t^{(n)}$ 下降。同时，短期债券的售卖将导致当前市场价格的下降和 r_t 的上升。因此，等式（17）会迅速（瞬间）趋于重新成立。

超过滚动投资序列的美元长期债券投资的预期“超额”或“异常”利润定义如下：

$$E_t(AP_t) \equiv R_t^{(n)} - E_t(r_{t+j}{}'s) \qquad (18)$$

其中，$E_t(r_{t+j}{}'s)$ 代表的是（17）式右边的部分。在 PEH 中应用的是即期利率，因此，上式意味着预期超额（或异常）利润等于零。与前面根据持有期收益率来对模型进行分类一样，我们也可以根据 $E_t(AP_t)$ 来对模型进行分类。表 1 对期限结构模型进行了总结，在此不再进行任何更多的评论。（17）两边取对数，使用近似公式 $\ln(1+z) = z$ 并让 z 趋于零，我们就可以得到如下近似线性关系

$$R_t^{(n)} = (1/n)[r_t + E_tr_{t+1} + E_tr_{t+2} + \cdots + E_tr_{t+n-1}] \qquad (19)$$

根据连续复利利率，类似分析可以得到如下关系

$$A\exp(R_t^{(n)}n) = E_t[A\exp(r_t)\exp(r_{t+1})\cdots\exp(r_{t+n-1})] \qquad (20)$$

$R_t^{(n)}$ 和 r_{t+i} 都是连续复利利率。两边同时取对数，我们有

$$R_t^{(n)} = (1/n)\ E_t\ [r_t + r_{t+1} + r_{t+2} + \cdots + r_{t+n-1}] \tag{21}$$

Jensen 不等式表明 $E\ [f\ (x)] \neq f\ (Ex)$，如果不考虑 Jensen 不等式，那么上式是一个准确的恒等关系。事实上，(21) 一般不是一个准确的恒等关系，因为 $E_t\ (\exp r_{t+j}) \neq \exp\ (E_t r_{t+j})$。然而，Campbell (1986) 认为，从实践的角度来看，我么可以不考虑 Jensen 不等式。

远期利率

远期利率可以用合适的即期利率来进行计算。例如，以 $R_t^{(2)}$ 进行的两年期投资所应得到收益，应该与以即期利率 r_t 进行一年期投资后在 t 和 $t+1$ 之间以远期利率 $f_{12,t}$ 再进行投资所获得的收益相等（参见 Cuthbertson 和 Nitzsche (2001a)）。因为在 t 时刻远期利率是已知的，所以以远期利率进行的投资属于无风险投资。利用连续复利（年）利率，有

$$2R_t^{(2)} = r_t + f_{12,t} \tag{22}$$

$$f_{12,t} = 2R_t^{(2)} - r_t \tag{23}$$

对 (23) 和 (21) 进行比较后就可以知道，EH 意味着

$$E_t r_{t+1} = f_{12,t}，\text{或等价地，} E_t r_{t+1} - r_t = (f_{12,t} - r_t) \tag{24}$$

因此远期利率是未来即期利率的无偏预测值。整个时期中的远期利率都可以计算出来。在理性预期假设下，也可以根据不同期限 $(m,\ n)$ 的远期利率回归结果来对 EH 进行检验

$$r_{t+m}^{(n-m)} - r_t = \alpha + \beta\ (f_{mn,t} - r_t)\ + \varepsilon_{t+m} \qquad (n > m) \tag{25}$$

在 EH 的零假设下，$\alpha = 0$，$\beta = 1$。

收益率曲线

PEH 是对（即期）收益率曲线进行分析的基础。例如，从 t 时刻来看，如果人们预期短期率上升（也就是说，对任意的 j，有 $E_t r_{t+j} > E_t r_{t+j-1}$），那么长期利率 $R_t^{(n)}$ 将大于当前的短期利率 r_t。收益率曲线（$R_t^{(n)}$ 关于期限的图像）将向上倾斜，因为 $R_t^{(n)} > R_t^{(n-1)} > \cdots > r_t$。相反的情形同样也成立，因此，当 $E_t r_{t+j} < E_t r_{t+j-1}$（对任意的 j）时，收益率曲线是向下倾斜的。因为通货膨胀预期会对将来的即期利率的预期产生影响（即 Fisher 效应），所以当未来几年的通货膨胀预期增加时，收益率曲线很有可能向上倾斜，如果确信中央银行采用的反通货膨胀策略，那么收益率曲线很有可能向下倾斜。如果存在具有单调性的流动性溢价，且该流动性溢价仅仅依赖于期限 n，$T^{(n)} > T^{(n-1)} > \cdots$，那么收益率曲线仍然和上面所描述的情形一样（Cuthbertson 和 Nitzsche (2001a)）。也可以在随机折现因子（SDF）模型框架中来分析收益率曲线，在这个框架中，风险溢价会对收益率曲线的形状产生影响。

20.4　小结

- 尽管财经媒体对到期收益率进行过广泛报道，它仍然是债券整个生命周期内年平均收益率的一个误导性测度指标，因为它假设息票能以当前的 YTM 进行再投资。
- 付息债券的单期 HPR 包括资本利得和息票支付两个部分。在有效金融市场上，债券价格可以视为用即期利率对将来的确定性息票支付（外加赎回价值）进行贴现后的现值。
- 期限结构各种理论的主要区别在于，期限溢价是否是：(a) 零；(b) 相对于时间和期限而言是常数；(c) 相对于时间而言是常数，但不同期限所对应的期限溢价不同；(d) 依赖于财富中“长期债务”的持有比例。这些假设分别对应于纯预期假说（PEH）、预期假说（EH）、流动偏好假说（LPH）和市场分割假说。可以用一种很特别的方式和更为正式的 SDF 方法将时变风险溢价嵌入到 EH 方程中去。
- 预期假说（EH）中所使用的是即期利率，这意味着长期利率是预期的未来短期利率的

加权平均。投资者是风险中性的，他们将整个投资期间中的预期收益率视为相等。因此，“长”“短”期债券之间的转换不可能获得异常利润，EMH 在风险中性下成立。

* 收益率曲线是债券的（即期）收益率 $R_t^{(n)}$（在某一个时点）关于期限 n 的图像。收益率曲线的形状主要由投资者对未来短期利率的预期来确定。如果预期未来短期利率上升（或下降），那么收益率曲线向上（或向下）倾斜。

第21章

预期理论：从理论到实证

目的

- 分析期限结构中预期假说（EH）（也就是说，假设期限溢价具有时间不变性或为零）的几种检验方法。
- 说明在预期假说（EH）中，预期超额单期 HPR（对任意期限的债券而言）不依赖于进行预测时所能获得的任何信息（有时称其为局部预期假说，但我们并不使用这一术语）。
- 证明在预期假说（EH）中，长期利率方程有如下含义：当前的收益率差（长期利率与短期利率之差）是下一期长期利率变化的最优预测。
- 说明预期假说也可以写成与未来短期利率方程相对应的形式，这意味着在长期债券的整个有效期内，收益率差是短期利率未来变化的最优预测。如果实际收益率差比短期利率在将来的预期变化高，那么这种情形就是过度反应假说。
- 在 Var 预测框架的基础上导出 EH 的检验方法。
- 说明如何应用 EH 来对收益率（或收益率差）进行方差边界检验（variance bounds test）。

经济金融学家对 EH 的检验是否一致、套利机会是否存在一直存在很大的争论（Cox、Ingersoll 和 Ross（1981），McCulloch（1993），Fisher 和 Gilles（1998））。Campbell（1986）认为任何差异对实证研究而言都不重要，而 Longstaff（2000b）认为，当市场不完备时，如果没有套利机会，那么 EH 的这些形式一般是成立的。因此，我们所进行的检验是针对 EH 版本的检验。注意，常数期限溢价使得理论模型或所使用的检验之间几乎不存在差异。关键的区别在于，我们是否考虑时变期限溢价。我们将在后面对其进行考察。

21.1　EH 的各种表示

EH 理论意味着在任何持有期内，不同期限债券的预期收益应相等。例如，持有年期零息债券一年所获得的收益（即资本利得）应等于已知的一年期无风险债券的收益。根据不同期限债券的即期收益率得出一个等价公式是可能的。例如，如果当前时刻 10 年期债券的收益率大于一年期债券的收益率，那么这并不意味着 10 年期债券比 10 年期间内的滚动投资能获得更多的收益。在这种情形中，PEH 意味着我们可以预期一年的收益率在将来（即下一年到第九年）会增加，因为一年期投资序列的预期收益等于 10 年期债券更高的收益率，因此，可以用多种等价方式来表示 EH，我们的一个任务就是指出这些等价表示之间的关系。

出于完整性考虑，以下推导包含了风险溢价。我们将说明如何用多种方式来表示 EH，每

一种方式都可以由如下版本的 EH 来进行推导：（连续复利）预期 HPR 等于无风险单期利率加上风险溢价，

$$E_t h_{t+1}^{(n)} \equiv E_t\ [\ln P_{t+1}^{(n-1)} - \ln P_t^{(n)}]\ = r_t + T_t^{(n)} \tag{1}$$

债券价格的对数为

$$\ln P_t^{(n)} = -R_t^{(n)} n \tag{2}$$

利用（1）和（2），我们立即可以得到如下式子：

长期利率变化方程（Change - in - long - rate equation）

$$(n-1)\ \{E_t R_{t+1}^{(n-1)} - R_t^{(n)}\}\ = \beta_L S_t^{(n)} + T_t^{(n)} \tag{3}$$

在 EH 下，$\beta_L = 1$；在 PEH 下，$T_t^{(n)}$ 等于零。方程（1），尽管以 HPR 为基础，还是给出了未来短期利率的期限结构关系。将 $S_t^{(n)} = R_t^{(n)} - r_t$ 代入（3）式，由 EH 可以得到如下向前差分方程（forward difference equation）

$$nR_t^{(n)} = (n-1)\ E_t R_{t+1}^{(n-1)} + r_t + T_t^{(n)} \tag{4}$$

（4）式滞后一期就得到

$$(n-1)\ R_{t+1}^{(n-1)} = (n-2)\ E_{t+1} R_{t+2}^{(n-2)} + r_{t+1} + T_{t+1}^{(n-1)} \tag{5}$$

利用 $E_t E_{t+1} = E_t$ 对（5）求数学期望并将结果代入（4）中，

$$nR_t^{(n)} = (n-2)\ E_t R_{t+2}^{(n-2)} + E_t\ (r_{t+1} + r_t)\ + E_t\ (T_{t+1}^{(n-1)} + T_t^{(n)}) \tag{6}$$

对（6）式右边第一项反复进行替代，注意对 $j = n$ 而言，$(n-j)\ E_t R_{t+j}^{(n-j)} = 0$，我们就得到如下结论：

长期利率是未来短期利率的加权平均

$$R_t^{(n)} = E_t R_t^{*(n)} + E_t \Phi_t^{(n)} \tag{7}$$

其中，

$$R_t^{*(n)} = \frac{1}{n} \sum_{i=0}^{n-1} r_{t+i} \quad \text{完全预见的长期利率} \tag{8a}$$

$$\Phi_t^{(n)} = \frac{1}{n} \sum_{i=0}^{n-1} T_{t+i}^{(n-1)} \quad \text{平均风险溢价} \tag{8b}$$

因此，n 期长期利率等于预期的未来短期利率 r_{t+i} 加上 n 期债券的平均风险溢价 $\Phi_t^{(n)}$。我们在前面已经导出过这一公式，现在可以看到，该公式和长期利率变化方程（3）一样是 HPR 方程（1）的替代性表示。变量 $R_t^{*(n)}$ 被称为完全预见长期利率（perfect foresight long rate），因为它是用单期短期利率 r_{t+i} 计算的加权平均“最终”价值。将 r_t 代入方程（7）的两边，我们就得到如下等价表示：

未来短期利率方程（Future - short - rate equation）

$$E_t S_t^{*(n,1)} = S_t^{(n,1)} + E_t \Phi_t^{(n)} \tag{9}$$

其中，

$$S_t^{(n,1)} = R_t^{(n)} - r_t \quad \text{实际收益率差} \tag{10a}$$

$$S_t^{*(n,1)} = \sum_{i=1}^{n-1} (1 - i/n)\ E_t \Delta r_{t+i} \quad \text{完全预见的收益率差} \tag{10b}$$

根据理性预期（RE），可以通过如下回归来对 EH 进行检验

$$S_t^{*(n,1)} = \alpha + \beta_S S_t^{(n,1)} + \gamma \Omega_t + v_{t+n-1} \tag{11}$$

其中，Ω_t 是由 t 时刻已知的变量所组成的一个向量，v_{t+n-1} 是 $n-1$ 个未来即期利率的 RE 预测误差（$r_{t+j} = E_t r_{t+j} + \eta_{t+j}$）的加权平均。在 EH 下，我们期望 $\beta_S = 1$，$\gamma = 0$。方程（9）说明，n 期利率和单期利率之间的实际收益率差 $S_t^{(n,1)}$ 等于未来短期利率预期变化的加权平均再加上期限溢价。变量 $S_t^{*(n,1)}$ 作为完全预见收益率差是已知的。正如我们将要看到的那样，方程（7）

中的变量（即 $R_t^{(n)}$ 和 r_{t+i}）通常是非平稳序列，而 $S_t^{(n,1)}$ 和 Δr_{t+i} 是 I（0）序列。因此，可以根据标准分布对方程（9）进行计量检验（而方程（7）则不能）。方程（7）和（9）是期限结构关系的一般表示，但它们却没有任何可操作性，除非我们假设期限溢价具有特殊的形式。

也许值得记住的是，当 $n=2m$ 时（如6月期债券和3月期债券），长期利率与短期利率之间的回归方程（3）和（9）是等价的。如果对所有（n，m）（3）成立，那么（9）也成立。然而，如果在（n，m）值的某个子集上（3）被拒绝，那么方程（9）不必再成立，因此方程（9）给出了EH有效性的一些不相关信息。在早期的实证研究中，对上述两个公式的实证基本上是 $n=2m$ 的情形，事实上，一般都是针对数据可获得的3月期和6月期纯贴现债券（如国库券）等进行实证研究。因此，两个回归方程在统计意义上是等价的。然而，其他期限情形（也就是说 $n \neq z$）也能获得一些结果，我们将在下一章中报告这些结果。

方差边界检验

“未来短期利率”方程（9）给出的是方差边界检验，它等价于在EH零假设下（常数期限溢价）的回归方程（11）。使用历史数据，我们能计算出完全预见收益率差 $S_t^{*(n,1)} = \sum_{i=1}^{n-1}\left(1-\frac{i}{n}\right)\Delta r_{t+i}$，并将其与实际收益率差 $S_t^{(n,1)}$ 进行比较。事实上，EH（常数期限溢价）意味着

$$\operatorname{Var}\left(S_t^{*(n,1)}\right) \geqslant \operatorname{Var}\left(S_t^{(n,1)}\right) \tag{12}$$

因为 $\operatorname{Var}(v_{t+n-1}) \geqslant 0$，且在RE下 $\operatorname{Cov}(S_t^{(n-1)}, v_{t+n-1})=0$。如果利率变化序列和收益率差序列都是平稳 I（0）的变量（这也是我们所想象的情形），那么利用收益率差序列所进行的标准方差边界检验是有效的。然而，从历史发展来看，这种类型的波动率检验最初却是基于利率水平值的检验（Shiller（1989））。完全预见的长期利率为

$$R_t^* = \frac{1}{n}\left(r_t + r_{t+1} + r_{t+2} + \cdots + r_{t+n-1}\right) \tag{13}$$

在EH＋RE下（即 $r_{t+i} = E_t r_{t+i} + \eta_{t+i}$），我们有

$$R_t^* = R_t + \omega_t \tag{14}$$

其中，$\omega_t = (1/n)\sum_{i=1}^{n-1}\eta_{t+i}$，$E_t(\eta_{t+i} \mid \Omega_t)=0$。使用历史数据，我们能构造时间序列 R_t^*。在RE下，投资者的预测是无偏的，因此随着 n 的增大，预测误差和 ω_t 趋于零。于是，对长期债券而言，我们预期完全预见的长期利率 R_t^* 在很宽的范围内与实际长期利率 R_t 相同。显然，在RE下，预测误差 ω_t 独立于 Ω_t，因此预测误差 ω_t 独立于 R_t。所以，根据（14），我们有

$$\operatorname{Var}(R_t^*) \geqslant \operatorname{Var}(R_t) \tag{15}$$

因为根据RE，$\operatorname{Cov}(R_t, \omega_t)=0$，$\operatorname{Var}(\omega_t) \geqslant 0$。同样，在EH的零假设下，

$$R_t^* = a + bR_t + c\Omega_t + \varepsilon_t \tag{16}$$

我们期望 $b=1$，$c=0$。“早期”的这些波动率检验在当时当然是一种创新，但是，如果用利率水平值的“粗糙”数据来进行实证研究，且序列不是平稳序列（很有可能不是平稳序列），那么所得出的方差是没有定义的，检验结果也值得怀疑。后来，学者们用收益率差 S_t 和完全预见的收益率差 S_t^* 来进行相同的检验，我们期望得到 $\operatorname{Var}(S_t^*) \geqslant \operatorname{Var}(S_t)$。

从方程（3）和（10）可以明显看出，如果我们忽略时变期限溢价（它与收益率差之间存在一定的相关性），那么收益率差的OLS回归系数不是一致估计。估计结果不一致的程度越高，相对于收益率差的方差而言，期限溢价的方差也就越大，两个变量之间的相关性也就越强。也有人（Walsh（1998））认为，即使EH成立，货币当局为了应对收益率差的变化也会改变短期利率，因此收益率差的回归系数也将不是一致估计。

方程（11）意味着，实际收益率差 $S_t^{(n)}$ 是未来短期利率预期变化 $E_t S_t^{*(n)}$ 的最优预测，但这一结论只是在期限溢价具有时间不变性的条件下才成立。预期单期期限溢价 $E_t T_t^{(n)}$ 与未来期

限溢价的平均值 $E_t\Phi_t^{(n)}$ 之间的差异在以下情形中可能很重要。

EH 中所使用的即期收益率假设投资者都是风险中性的，预期收益是投资者投资决策的唯一根据。收益的内在变化或不确定性对投资决策不会有任何影响。根据我们的方程，EH 意味着对所有的 n，有 $T^{(n)} = \Phi^{(n)} = 0$。

在计算完全预见收益率差 S_t^* 的过程中，如果不用最终值 r_{t+j} 去替换 $E_t r_{t+j}$，那么我们就可以用显式方程来预测 r_{t+j}。正如我们即将看到的那样，这种预测结果被称为理论收益率差（theorectical spread），记为 S_t'。现在，我们已有好几种方法来（如下所述）对 EH 进行检验，在零假设下这些检验方法都是等价的。

- 超额 HPR 回归；
- 长期利率变化回归；
- 未来短期利率变化回归，使用最终值——完全预见的收益率差；
- 未来短期利率变化回归，使用显式预测方程——理论收益率差；
- 方差边界检验。

21.2 Var 方法

在这一节中，我们讨论如下问题：

- 在 Var 预测方程中，EH 如何对交叉方程中的参数进行限制。
- 如何应用 Var 来对未来短期利率（理论收益率差 S_t'）进行预测，然后用方差比统计量和相关系数来对理论收益率差 S_t' 和实际收益率差 S_t 进行比较。

在对未来短期利率进行多期预测时，我们可以使用 Campbell - Shiller（C - S）向量自回归（Var）方法。我们用一个简单例子来说明这种方法。在三个时期中应用 EH 就可以得到

$$R_t = \frac{1}{3}\left(r_t + E_t r_{t+1} + E_t r_{t+2}\right) \tag{17}$$

对上式重新进行参数化处理后就得到

$$S_t = \frac{2}{3}\Delta E_t r_{t+1} + \frac{1}{3}\Delta E_t r_{t+2} \tag{18}$$

其中，$S_t = R_t - r_t$ 为长短期收益率差，$E_t\Delta r_{t+1} = (E_t r_{t+1} - r_t)$，$E_t\Delta r_{t+2} = (E_t r_{t+2} - E_t r_{t+1})$。现在，假设 S_t 和 Δr_t 都可以表示为一阶（为了简化起见）双变量向量自回归（Var）的形式：

$$S_{t+1} = a_{11}S_t + a_{12}\Delta r_t + \omega_{1,t+1} \tag{19a}$$

$$\Delta r_{t+1} = a_{21}S_t + a_{22}\Delta r_t + \omega_{2,t+1} \tag{19b}$$

或用向量表示为

$$z_{t+1} = Az_t + \omega_{t+1} \tag{20}$$

其中，$z_{t+1} = (S_{t+1},\ \Delta r_{t+1})'$，$A$ 是系数 a_{ij} 所构成的（2×2）矩阵，$\omega_{t+1} = (\omega_{1,t+1},\ \omega_{2,t+1})'$。根据（20），使用预测的链式法则就得到未来 z 的最优预测

$$E_t z_{t+1} = Az_t \tag{21a}$$

$$E_t z_{t+2} = E_t A z_{t+1} = A^2 z_t \tag{21b}$$

现在，令 $e1' = (1,\ 0)$，$e2' = (0,\ 1)$ 为 2×1 选择变量。于是

$$S_t = e1'z_t \tag{22a}$$

$$E_t\Delta r_{t+1} = e2'E_t z_{t+1} = e2'Az_t \tag{22b}$$

$$E_t\Delta r_{t+2} = e2'E_t z_{t+2} = e2'A^2 z_t \tag{22c}$$

将以上方程代入 PEH 方程（18）就得到

$$S_t = e1'z_t = \left(\frac{2}{3}e2'A + \frac{1}{3}e2'A^2\right)z_t \tag{23}$$

Var 中的交叉方程限制

如果（23）对所有 z_t 都成立，那么系数 a_{ij} 之间必定存在如下非线性关系：

$$f(a) = e1' - e2'\left(\frac{2}{3}A + \frac{1}{3}A^2\right) = 0 \tag{24}$$

其中，$f(a)$ 定义为限制集合。EH 的检验和用 Var 表示的预测方法需要对无限制 Var 方法进行估计，需要在限制（24）的基础上进行统计检验。

WALD 检验和似然比检验

我们在这里对 Wald 检验的形式进行简略叙述是值得的。在对（2×2）Var 进行估计后，我们就得到无限制 Var 系统方差协方差矩阵的估计结果：

$$\sum = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} \tag{25}$$

如果不存在异方差，那么误差项的方差和协方差可以由每个方程的残差来进行计算（如 $\sigma_1^2 = \sum_t \hat{\omega}_{1t}^2/n$，$\sigma_{12} = \sum_t \hat{\omega}_{1t}\hat{\omega}_{2t}/n$）；除此之外，我们得用 GMM 对异方差进行纠正。（24）中非线性函数 $f(a)$ 的方差协方差矩阵为：

$$\text{Var}[f(a)] = f_a(a)' \sum f_a(a) \tag{26a}$$

其中，$f_a(a)$ 是限制条件关于参数 a_{ij} 的导数。*Wald* 检验统计量为

$$W = f(a)\{\text{Var}[f(a)]\}^{-1} f(a)' \tag{26b}$$

我们很难从 Wald 检验的一般形式中获得一些直觉（参见 Buse（1982））。然而，$f(a)$ 的方差越大，W 的值就越小。因此，A 矩阵的估计越不准确，W 的值就越小，通过 Wald 检验的可能性就越大（即不拒绝零假设）。另外，如果限制条件成立，那么 $f(a) \approx 0$，$W \approx 0$。可以证明，在误差项的标准假设下，W 在零假设下的渐进分布是自由度为 r 的 χ^2 分布，r 等于限制条件的个数。如果 W 小于临界值 χ_c^2，那么我们就不能拒绝零假设 $f(a) = 0$。

从原理上讲，我们也能用似然比检验方法。似然比检验要对无限制条件的两个方程组成的系统和带限制条件的系统进行比较。记受限 Var 的方差协方差矩阵为 Σ_u，记无限制 Var 的方差协方差矩阵为 Σ_r，这里所施加的限制方程为（24）。似然比检验统计量的计算公式为

$$LR = n.\ln[(\det \sum\nolimits_r)/(\det \sum\nolimits_u)] \tag{27}$$

其中，n = 样本观察值的数目，"det" 表示协方差矩阵的行列式值，即 $\det \sum = \sigma_1^2\sigma_2^2 - (\sigma_{12})^2$。如果数据表明限制条件成立，那么我们就不能期望残差有太大的变化，因此 $\det \sum_r \approx \det \sum_u$，从而 LR≈0。反之，如果数据与限制条件不相容，那么我们预期的拟合效果应该不是很好，限制方程的残差要比（等价的）无限制方程的残差要大（平均而言），因此 $\det \sum_r / \det \sum_u$ 很大，LR 也很大。在零假设下，LR 的渐进分布是自由度为 q 的 χ^2 分布，这里的 q 是限制参数的个数。因此，如果 LR > $\chi_c^2(q)$，那么我们就拒绝零假设，这里的 $\chi_c^2(q)$ 是临界值。不幸的是，*LR* 检验要求在更为复杂的交叉方程限制条件下来对 Var 进行估计。显然，这种交叉方程限制在这里并不是很直观（参见 Bekaert，Hodrick 和 Marshall（1997））。因此，尽管 Wald 检验存在一些缺陷，在一般情况下，我们还是使用 Wald 检验。

Var - Wald 检验程序的使用非常普遍，可以用它来对更为复杂的期限结构关系进行考察究，应用 Var 的高阶滞后项就可以完成相应的检验。Campbell 和 Shiller（1991）认为，在 EH 下，n 期和 m 期（$n > m$）债券收益率之差 $s_t^{(n,m)}$ 可以表示为

$$s_t^{(n,m)} = E_t \sum_{i=1}^{k-1} (1 - i/k)\ \Delta^m r_{t+im}^{(m)} \tag{28}$$

其中，$\Delta^m r_t = r_t - r_{t-m}$，$k = n/m$（整数）。例如，当 $n = 4$，$m = 1$ 时，

$$S_t^{(4,1)} = R_t^{(4)} - r_t^{(1)}$$

$$S^{(4,1)} = E_t\left(\frac{3}{4}\Delta r_{t+1}^{(1)} + \frac{2}{4}\Delta r_{t+2}^{(1)} + \frac{1}{4}\Delta r_{t+3}^{(1)}\right) \tag{29}$$

高阶 Var 的形式为

$$\begin{aligned} S_{t+1} = {} & (a_{11}S_t + a_{12}\Delta r_t) + (a_{13}S_{t-1} + a_{14}\Delta r_{t-1}) \\ & + (a_{15}S_{t-2} + a_{16}\Delta r_{t-2}) + \cdots + \omega_{1,t+1} \end{aligned} \tag{30a}$$

$$\begin{aligned} \Delta r_{t+1} = {} & (a_{21}S_t + a_{22}\Delta r_t) + (a_{23}S_{t-1} + a_{24}\Delta r_{t-1}) \\ & + (a_{25}S_{t-2} + a_{26}\Delta r_{t-2}) + \cdots + \omega_{2,t+1} \end{aligned} \tag{30b}$$

然而，如果按照通常方式已经得到 a_{ij} 的估计值，那么我们可以将“高阶”系统重新排列为一阶系统。例如，假设 Var 的阶 $p=2$，那么用矩阵符号表示的等价系统为

$$\begin{bmatrix} S_{t+1} \\ \Delta r_{t+1} \\ S_t \\ \Delta r_t \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} S_t \\ \Delta r_t \\ S_{t-1} \\ \Delta r_{t-1} \end{bmatrix} \begin{bmatrix} \omega_{1,t+1} \\ \omega_{2,t+1} \\ 0 \\ 0 \end{bmatrix} \tag{31}$$

方程（31）是 Var 的伴随形式（campanion form），它有如下简写形式

$$Z_t = AZ_{t-1} + \omega_{t+1} \tag{32}$$

其中，$Z'_{t+1} = [S_{t+1}, \Delta r_{t+1}, S_t, \Delta r_t]$。给定 $(2p\times 1)$ 选择向量 $e1' = [1, 0, 0, 0]$，$e2' = [0, 1, 0, 0]$，我们有

$$S_t^{(n,m)} = e1'Z_t \tag{33a}$$

$$\Delta r_t^{(m)} = e2'Z_t \tag{33b}$$

$$E_t\Delta r_{t+j}^{(m)} = e2'A^jZ_t \tag{33c}$$

在我们的例子中，$n=4$，$m=1$，$p=2$。如果将（33）代入一般的短期利率方程（9），那么 Var 的非线性限制可以表示为

$$f(a) = e1' - e2'A\left[I - \frac{m}{n}(I - A^n)(I - A^m)^{-1}\right](I - A)^{-1} = 0 \tag{34}$$

在我们的例子中，相应的表示是

$$f^*(a) = e1' - e2'A\left[I - \frac{1}{4}(I - A^4)(I - A)^{-1}\right](I - A)^{-1} = 0 \tag{35}$$

交叉方差限制的解释

我们回到三个时期中的 EH 情形，考察从（2×2）Var 非线性限制条件中是否能获得一些直觉。

$$E_tS_{t+1} = a_{11}S_t + a_{12}\Delta r_t \tag{36a}$$

$$E_t(\Delta r_{t+2}) = a_{21}E_tS_{t+1} + a_{22}E_t\Delta r_{t+1} \tag{36b}$$

应用上述方程和（19b），

$$\begin{aligned} E_t(\Delta r_{t+2}) &= a_{21}(a_{11}S_t + a_{12}\Delta r_t) + a_{22}(a_{21}S_t + a_{22}\Delta r_t) \\ &= (a_{21}a_{12} + a_{22}^2)\Delta r_t + a_{21}(a_{11} + a_{22})S_t \end{aligned} \tag{37}$$

在 EH 方程（18）中应用（19b）和（37）就得到，

$$\begin{aligned} S_t &= \frac{2}{3}(a_{21}S_t + a_{22}\Delta r_t) + \frac{1}{3}[(a_{21}a_{12} + a_{22}^2)\Delta r_t + a_{21}(a_{11} + a_{22})S_t] \\ &= \left[\frac{2}{3}a_{22} + \frac{1}{3}(a_{21}a_{12} + a_{22}^2)\right]\Delta r_t + \left[\frac{2}{3}a_{21} + \frac{1}{3}a_{21}(a_{11} + a_{22})\right]S_t \\ &= f_1(a)\Delta r_t + f_2(a)S_t \end{aligned} \tag{38}$$

令方程（38）两边的系数相等，非线性限制条件可以表示为

$$0 = f_1(a) = \frac{2}{3}a_{21} + \frac{1}{3}(a_{21}^2 + a_{22}a_{11}) \tag{39a}$$

$$0=f_2\ (a)\ =\frac{2}{3}a_{22}+\frac{1}{3}a_{22}\ (a_{21}+a_{12}) \tag{39b}$$

用其他方法来导出这些条件是一件十分乏味的事情，相对而言，矩阵形式要容易得多。当 $n=3$，$m=1$ 时，矩阵限制条件（34）必然与（39）等价，证明过程作为一道简单的习题留给读者。与前面的情形一样，非线性交叉方程限制（39）可以保证：

(ⅰ) 基于 Var 中信息集合 Λ_t 的预期（异常）利润为零；

(ⅱ) EH 方程（18）成立；

(ⅲ) Var 的预测误差 Δr_{t+1} 和 Δr_{t+2} 与 t 时刻或之前的信息独立。这就是 RE 的正交性质。

理论收益率差 $S_t^{'}$ 和实际收益率差 S_t

Campbell 和 Shiller（1991）建议使用其他方法来检验 EH 理论是否成立，就三个时期（即 $n=3$，$m=1$）的情形而言，我们将其概述如下：

$$S_t=\frac{2}{3}E_t\Delta r_{t+1}+\frac{1}{3}E_t\Delta r_{t+2} \tag{40}$$

方程（40）的右边可以用 Var 的线性预测值来进行度量，理论收益率差 $S_t^{'}$ 为

$$S_t^{'}=e2'\left[\frac{2}{3}A+\frac{1}{3}A^2\right]z_t=f\ (A)\ z_t=f_1\ (a)\ \Delta r_t+f_2\ (a)\ S_t \tag{41}$$

理论收益率差是计量经济学家的最佳预测（best shot），短期利率未来变化的真实（RE 下）预测（加权平均）就是这个预测值。如果 EH（40）成立，那么 $f_2\ (a)\ =1$，$f_1\ (a)\ =0$，因此 $S_t^{'}=S_t$，实际收益率差与理论收益率差高度相关。在实际数据集合中，后一个限制条件通常不会精确成立，于是我们预期方程（41）中的 $S_t^{'}$ 与实际的 S_t 有较大的差异。因此，在 PEH + RE 的零假设下，如下“统计量”是 PEH 是否成立的有用度量指标（参见 Engsted（2002））。

* 在 S_t 和 $S_t^{'}$ 关于时间的图像中，两个序列在很宽范围内的运动是一致的，但 S_t 的峰值比 $S_t^{'}$ 的峰值大。

* 在回归方程 $S_t^{'}=\alpha+\beta S_t+v_t$ 中，我们期望 $\alpha=0$ 和 $\beta=1$。

* S_t 和 $S_t^{'}$ 的相关系数等于 1，与此类似，要么方差率比 VR 等于 1，要么标准差比率 SDR 等于 1。

$$\begin{gathered}\mathrm{VR}=\mathrm{Var}\ (S_t)\ /\mathrm{Var}\ (S_t^{'}) \\ SDR=\sigma\ (S_t)\ /\sigma\ (S_t^{'})\quad \text{且}\quad \mathrm{Corr}\ (S_t,\ S_t^{'})\ =1\end{gathered} \tag{42}$$

我们有时会给 SDR 一个非常宽松的名字“方差比率”，因为两者给出的是相似的推断。过度反应假设意味着 SDR > 1 或 $\beta>1$，因为实际收益率差的运动比未来短期利率的预期值（加权平均）（即理论收益率差 $S_t^{'}$）要大。

EH 方程（40）意味着 S_t 是短期利率未来变化的一个充分统计量，因此 S_t 是利率变化的“Granger 原因”。这意味着，在 Var 方程中就 Δr_{t+1} 所对应的（19*b*）而言，S_t 的滞后值（作为一个组）应部分地有助于解释 Δr_{t+1}（这里要用到所谓的块外生性检验（block exogeneity test））。

假设 R_t 和 r_t 都是 I（1）变量，因此 Δr_{t+j} 是 I（0）序列。如果 EH 正确，那么由（18）可知，收益率差 $S_t=R_t-r_t$ 也必须是 I（0）序列。因此，R_t 和 r_t 必须是协整的，协整参数为 1。

值得注意的是，如果计量经济学家认为“准确”的 RE 模型（40）成立，那么 S_t 和 $S_t^{'}$ 在所有时期中都应相等。后一观点并不意味着理性投资者就不会有预测误差，他们确实可能产生预测误差。然而，S_t 在市场上是指未来利率的预期值。但是，如果 RE 方程“不正确”，那么 $S_t-S_t^{'}$ 是这一随机偏差的度量指标（即（40）中的误差项）。

就完全预见收益率差而言，Var 方法的优点在于能避免重叠数据（overlapping data）的使用，而在小样本情形中使用重叠数据所导致的误差修正（使用 GMM 方法）可能是有问题的。Var 的缺点在于需要（S_t，Δr_t）的显性预测形式，如果假设错误，那么估计结果将是有偏的。

Wald 检验也具有很差的小样本性质，根据非线性限制构造的准确方式，它不具有不变性（Gregory 和 Veall（1985））。Wald 检验也可能拒绝 EH + RE 的零假设，因为零假设下数据存在“细微偏差”。例如，如果（39）中 $f_1(a)=1.03$ 但 $f_1(a)$ 的标准误差为 0.003，那么我们就应拒绝 $H_0: f_1(a)=1$（t 检验），但经济学家仍然会说数据在很大程度上是支持 EH 的。

一句警告的话是：不要将 Var 方法中的完全预见收益率差 S_t^* 和理论收益率差 S_t' 相混淆。完全预见收益率差 S_t^* 是用实际值来代替预期变量之后所构造的一个变量，它是投资者在 100% 预测精度下的收益率差。另一方面，理论收益率差是未来利率的一种显式预测，这种预测是建立在计量经济学家对一些变量猜测的基础上，而这些变量是投资者在对未来收益率进行预测时实际使用的变量。

21.3 时变期限溢价——Var 方法

在考虑时变期限溢价时，EH 的检验要求 Var 模型包含三个变量：不仅包含收益率差和短期利率变化两个变量（像上面所使用的那样），还包括持有期收益率变量。后一个变量（参见方程（1））可以捕获（平稳的）期限溢价的一些运动特征（有一定的误差）（Tzavalis 和 Wickens（1997））。考虑包含如下向量的三变量 Var 系统：

$$Z_t=[S_t,\ \Delta r_t,\ h_t-r_{t-1}] \tag{43}$$

向量中的变量都是平稳变量，因此多变量 Wold 表示存在（Hannan（1970）），该表示可以用 p 阶 Var 来逼近，表示有如下伴随形式

$$Z_{t+1}=AZ_t+v_{t+1} \tag{44}$$

方程（1）意味着，只有在期限溢价具有时间不变性的条件下预期超额持有期收益 $E_t h_{t+1}-r_t$ 才是一个常数，根据 Var 模型，这意味着（因为所有变量都表示为“与均值的偏差”的形式）

$$e3'A=0 \tag{45}$$

与两变量 Var 模型相比，这是唯一的新增限制条件，短期利率方程意味着其他限制条件（34）、（39a）~（39b）在三变量情形中仍然成立。如果（39a）~（39b）遭到拒绝，那么未来平均期限溢价的预期 $\Phi_t^{(n)}=(1/n)\sum T_{t-i}^{(n-i)}$ 具有时变特征（参见（9））。

未来短期利率和期限溢价的意外变化

现在，我们处于对如下事件进行考察的位置：单期持有期收益 $eh_{t+1}=h_{t+1}-E_t h_{t+1}$ 方面的“意外变化”或“消息”在多大程度上要归因于未来利率方面的“消息”或期限溢价 $T_t^{(n)}$ 未来变化方面的“消息”。因此，我们要说明后者在实证研究中对拒绝 PEH（即假设常数或零期限溢价）的解释有多重要。

持有期收益率预料不到的变化必然归因于长期利率预料不到的变化（参见方程（1）和（2））。长期利率的变化要么是因为对未来短期利率预期进行了修正，要么是因为对未来（短期）期限溢价进行了修正（参见（7））。根据方程（1）~（3），我们就得到（参见 Tzavalis 和 Wickens（1997））

持有期收益率方面的“意外变化”= 未来利率方面的“消息”+ 未来期限溢价方面的“消息”

$$eh(n,\ t+1)=-er(n,\ t+1)-eT(n,\ t+1) \tag{46}$$

其中，

$$eh_{t+1}^{(n)}\equiv h_{t+1}^{(n)}-E_t h_{t+1}^{(n)}=[h_{t+1}^{(n)}-r_t]-[E_t h_{t+1}^{(n)}-r_t] \tag{47a}$$

$$er_{t+1}^{(n)}=(E_{t+1}-E_t)\sum_{i=1}^{n-1}\sum_{j=1}^{i}\Delta r_{t+j} \tag{47b}$$

$$eT_{t+1}^{(n)} = (E_{t+1} - E_t) \sum_{i=1}^{n-1} T_{t+i}^{(n-i)} \tag{47c}$$

方程（46）的直觉如下：对 n 期债券而言，如果单期收益率（$h_{t+1} - E_t h_{t+1}$）出现预料不到的上升，那么这必然归因于长期利率 $R_t^{(n)}$ 预料不到的下降，$R_t^{(n)}$ 预料不到的下降又反过来归因于当前或未来短期利率（即方程（47b）中的 $er_{t+1}^{(n)}$ 项）预料不到的下降。与之相反的是，h_{t+1} 预料不到的上升可能是因为投资者感知到未来风险溢价的平均值（即方程（47c）中的 $eT_{t+1}^{(n)}$ 项）会出现预料不到的下降。

在实证研究中，我们怎样考察未来期限溢价预期变化的相对重要性呢？根据方程（47a）~（47c），由 Var 方程（44）我们可以看到，“意外变化”依赖于 Var 的残差。特别地，$v_{3,t+1} = e3'v_{t+1}$ 是 Var 系统中第三个方程的误差项，代表超额持有期收益 $h_{t+1} - E_t h_{t+1}$ 的意外变化部分。同理，

$$\begin{aligned} eT_{t+1}^{(n)} &= -er_{t+1}^{(n)} - eh_{t+1}^{(n)} \\ &= e2' [(n-1) I + (n-2) A + (n-3) A^2 \\ &\quad + \cdots + (n-(n-1)) A^{n-2}] \times v_{2,t+1} - v_{3,t+1} \end{aligned} \tag{48}$$

第一项纯粹是未来短期利率意外变化部分的加权和（即 $(E_{t+1} - E_t) \sum_{i=1}^{n-1} \sum_{j=1}^{i} \Delta r_{t+j}$），$e2'$“选出了”误差项 v_{t+1} 的第二个元素，A 矩阵代表未来短期利率消息的持久程度。根据（46），如果未来期限溢价方面的消息非常少（即 $eT_{t=1}^{(n)} \approx 0$），那么单期收益 $eh_{t+1}^{(n)}$ 的意外变化部分完全反映了未来短期利率中的“消息”，因此 $eh_{t+1} = -er_{t+1}$，如下关系式也成立：

$$\sigma(er, eh) = 1 \tag{49a}$$

$$\rho(er, eh) = -1 \tag{49b}$$

另外，由方程（46）、（47a）和（1），我们就得到

$$h_{t+1}^{(n)} - r_t = T_t^{(n)} - er_{t+1}^{(n)} - eT_{t+1}^{(n)} \tag{50}$$

根据期限溢价 $T_t^{(n)}$ 对 Var 中变量充分替代的程度，单期持有期收益率回归方程（1）的残差反映了短期利率消息和期限溢价消息的影响。如果 $eT_{t+1}^{(n)} \approx 0$，那么 Var 中超额收益率方程（即第三个方程）的 $1 - R^2$ 反映的是超额持有期收益率由未来短期利率方面的消息所能解释的比例。

21.4　小结

* EH 意味着，n 期（即期）收益率 $R_t^{(n)}$ 和 m 期即期收益率 $R_t^{(m)}$ 之差 $S_t^{(n,m)}$ 是长期利率的预期未来变化和短期利率的预期未来变化两者的最优预测值，$S_t^{(n,m)}$ 导致了如下回归检验。

长期利率方程

$$((n-m)/m) [R_{t+m}^{(n-m)} - R_t^{(m)}] = \alpha + \beta_L S_t^{(n,m)} + \gamma \Omega_t + \varepsilon_{t+m}$$

未来短期利率方程

$$S_t^{(n,m)*} = \alpha + \beta_S S_t^{(n,m)} + \gamma \Omega_t + \eta_t$$

其中，完全预见收益率差为 $S_t^{(n,m)*} = (1/k) \sum_{i=0}^{k-1} \Delta^m R_{t+im}^{(m)}$，$k = n/m$ 为整数。在 EH + RE 下，我们预期 $\beta_S = \beta_L = 1$，$\gamma = 0$。非零常数期限溢价意味着 $\alpha \neq 0$。

* 未来短期利率方程可以通过如下方式来进行检验：使用 Δr_{t+j}，使用单方程回归来构造完全预见收益率差 S_j^*；或使用 Δr_{t+j} 的显式预测方程，使用理论收益率差 S_t'。后者就是 Var 方法。

* 在 Var 方法中，EH 意味着 $S_t' = S_t$，因此 $\sigma(S_t') = \sigma(S_t)$，$\mathrm{Corr}(S_t', S_t) = 1$。EH 也意味着 Var 参数存在交叉方程限制。

* 在 Var 中使用超额 HPR，我们能获得时变期限溢价行为的一些直觉。

* 在 EH 的检验中最早使用的是方差边界不等式，但现已基本上被上述以回归分析为基础的检验所取代。

第22章

期限结构的经验证据

目的

* 考察期限结构模型检验中所用的数据类型和质量。
* 考察长期利率和短期利率的性质及协整问题。
* 分析方差边界检验的早期工作，对长期利率和完全预见收益率差的时间序列行为进行比较。
* 陈述远期利率是否是未来即期利率无偏预测的证据。
* 陈述应用“长期利率变化”方程和“短期利率变化”方程来对预期假说（EH）进行单方程检验的结果。
* 用 Var 方法来给出 EH 各种检验的详细案例研究，包括时变期限溢价的可能性。
* 分析零假设下 EH 相互等价的各种检验为什么在实践中给出的是不同的推断。

在这一章里，我们用方差边界法、单方程回归法和 Var 方法来对 EH 进行检验。在本章的绝大部分内容中，我们集中关注具有时间不变性期限溢价的 EH 检验。在本章快要结束的部分里，我们考察那些明显有助于解释时变期限溢价的模型（但这不是基于无套利的 SDF 方法），我们将在第 23 章中考察这个问题。

22.1　数据与协整分析

实证研究使用数据的“质量”存在很大的差异，在不同研究之间进行比较会变得非常危险。研究人员使用的收益率数据代表的是开盘利率或收盘利率（月中某个特定日期），可以是“竞买利率”或“竞卖利率”，也可以是两者的平均。紧接着的一个问题与“时间”有关。如果我们试图对三年期债券滚动投资所获得收益和在三个一年期债券上的收益进行比较，那么 t 时刻长期债券收益率数据的测度时间应与短期利率完全一致，投资期限也应该完全吻合。换句话说，利率代表的是投资者采取投资策略时的实际（即期）利率。

正如我们在前面章节中所注意到的那样，根据原理，预期假说的检验需要即期收益率数据。对期限超过两年的债券而言，即期收益率数据无法直接获得，只能根据付息债券的价格来进行估计：这就涉及更为深入的逼近技巧。

利率的协整性质

假设长期利率 R_t 和短期利率 r_t 是 I（1）序列，根据构造，ΔR 和 Δr 是 I（0）序列。“未来短期利率”或“长期利率”方程（参见前面的章节）意味着对 PEH + RE 的弱检验 R_t 是 r_t

和 r_t 相互协整 1，且协整参数为。另外一种检验方法是：EH 意味着收益率差序列 $S_t = R_t - r_t$ 是 I（0）序列。尽管人们经常发现两种利率之间存在协整关系，收益率差序列 $S_t^{(n,m)}$ 也是平稳的，我们还是可以用更为全面的方式来进行协整检验。如果我们有 q 个 I（1）利率，那么 EH 意味着存在（$q-1$）个相互协整的收益率差。我们对单期利率 $R^{(1)} \equiv r$ 进行正规化处理，因此，对 $X_t = \{r, R^{(2)}, \cdots, R^{(q)}\}$ 而言，EH 意味着限制后的协整向量具有 $\{1, -1, 0, \cdots, 0\}$，$\{1, 0, -1, 0, \cdots, 0\}$ 等形式。（$q-1$）个收益率差序列中的一部分也会进入解释利率变化 ΔX_t 的向量 ECM 模型。Johansen（1988）的程序允许在如下形式的向量误差修正模型（VECM）中同时对这些协整向量进行检验：

$$\Delta X_t = \theta\ (L)\ \Delta X_{t-1} + \gamma' X_{t-1} + \varepsilon_t \tag{1}$$

其中，γ 是（$q \times q$）参数矩阵，$R = \{R^{(2)}, \cdots, R^{(q)}\}$

$$\begin{aligned}\Delta X_t = \theta\ (L)\ \Delta X_{t-1} &+ \alpha_1'\ (R^{(1)} - \beta_1' R)_{t-1} + \alpha_2'\ (R^{(1)} - \beta_2' R)_{t-1} \\ &+ \cdots + \alpha_{q-1}'\ (R^{(1)} - \beta_q' R)_{t-1} + \varepsilon_t\end{aligned}$$

首先，我们检验系统中协整向量的数目是否等于 $q-1$，接着检验 EH 隐含的协整参数限制的联合零假设。两者都是 EH 成立的“最小”检验。Shea（1992）和 Hall、Anderson 和 Granger（1992）发现，尽管使用美国数据我们也不能拒绝存在个协整向量的零假设，然而经常出现的情形是，数据拒绝 $\{1, -1, 0, \cdots, 0\}$，$\{1, 0, -1, 0, \cdots, 0\}$ 等类型的限制。抛开细微的统计问题不论（实证中存在很多细微的统计问题，但这些不仅是小样本所带来的问题），解释这些结果的一个关键考虑是，VECM 是否是利率数据生成过程的一种充分表示（如参数在整个样本期中是否为常数?）。如果 Var 是可接受的，但从数字（和统计）来看，一些限制条件与 $\{1, -1, 0, \cdots, 0\}$，$\{1, 0, -1, 0, \cdots, 0\}$ 等相去甚远，那么在整个期限范围内 EH 应该被拒绝。然而，在收益率数据集合的一些子集上（如较短的时期末，参见 Cuthbertson（1996），Cuthbertson、Hays 和 Nitzsche（1998））或数据集合的某些特殊子集上（参见 Hall、Anderson 和 Granger（1992）），EH 仍然可能成立。

最近，Hansen（2002）在已知时刻不存在结构性变化的假设下使用美国数据对 EH 所隐含的协整限制条件进行了检验，结果发现，在利率出现戏剧性上升且波动剧烈的 1979—1982 货币基础控制时期中存在明显的结构变化。在动态 VECM 模型考虑结构变化后就会发现存在满足 EH 的协整向量。除了进行结构变化检验之外，Soe（2003）可以解释为放弃线性 VECM、使用非线性阈值模型（TAR）来“消除 1979—1982 期间的数据问题”。在线性 VECM 中，协整参数已经非常接近 EH 中的 $\{1, -1, 0, \cdots, 0\}$，$\{1, 0, -1, 0, \cdots, 0\}$ 等。TAR 模型中出现的仍然是这种情形。但后者允许 R_t 或 r_t 的变化出现不同的动态反应，这取决于收益率差 $R_{t-1} - r_{t-1}$ 是位于上临界值之上、下临界值之下还是两个临界值之间。就美国而言，在 1973—1995 年期间，与线性 VECM 模型相比，TAR 模型有助于协整 VECM 模型解释力的增强。Sarno 和 Thornton（2003）使用美国 1974—1999 年期间联邦基金（FF）利率和国债（TB）利率的每日数据，运用双变量非线性不对称 VECM 模型进行了与 Soe（2003）类似的研究。协整向量（在整个样本中）为 $z = FF - 1.15TB + 0.5$，非均衡项分别以正负偏差的形式出现。另外，z 中的 ESTAR 项可以解释作用在 ΔF 或 ΔTB 上的对称非线性效应。这些非线性效应将残差减少了约 20%（也就是说，与线性 *VECM* 相比，回归标准差大约减小了 10%）。绝大多数调整是通过联邦基金利率进行的，正如作者所注意到的那样，这让人感到有些奇怪。因为联邦储备系统的目标是联邦基金利率，所以我们可以预期，为了与非均衡的相适应，短期国债利率要调整得更多一些。如果投资者能准确地预测联邦基金利率的政策变化，能在联邦基金利率变化之前改变短期国债利率（预测），那么数据就能说明联邦基金利率对这种非均衡有很强的滞后效应（但这要求市场能准确预测政策变化，这似乎是不可能的事情）。对 1979—1982 年货币目标时期中由于利率波动所导致的非线性效应进行考察是一件很有趣的事情。（我们在下面的部分中将看到，美国 1979—1982 年期间的数据对 EH 不同类型的检验有影响，但使用简单的虚拟变量就可以很容易地处理这些问题，效果比较令人满意）。上述结果可以总结为：

（ⅰ）就绝大部分时期而言，长期利率和短期利率是相互协整的，协整参数接近于1或等于1。

（ⅱ）就绝大多数期限而言，收益率差确实是利率未来变化的Granger原因（即方程（1）中有些α_i在统计上显著）。

（ⅲ）美国的利率之间存在一种动态非线性关系，但人们仍然不清楚对不同样本而言这种关系的稳健程度到底有多高，也不清楚这种关系与线性模型的估计结果之间到底有多大偏差。除了对“利率是相互协整的”和“协整向量具有{1，−1，0，…，0}、{1，0，−1，0，…，0}等形式”进行弱检验外，这些模型无法对EH进行仔细检验（Balke和Fomby（1997））。

无论有限样本中的协整检验结果如何，不能将收益率差视为平稳序列的观点意味着，随着期限的增加，（无条件）方差接近于无穷大，或等价地说，随着时间的变化，长期利率和短期最终会分离。只有勇敢的经济学家持有这种观点，因为这意味着风险溢价也是非平稳序列，从而是无界序列。

22.2　方差边界检验

如果EH＋RE成立，期限溢价只依赖于n（即具有时间不变性），那么实际长期利率的方差要小于完全预见长期利率的方差。因此，方差比

$$\mathrm{VR}=\frac{\mathrm{Var}\left(R_t^{(n)}\right)}{\mathrm{Var}\left(R_t^{(n)*}\right)} \tag{2}$$

小于1。然而，在早期的方差边界检验中，Shiller（1979）使用的是美国和英国1966—1977年的到期收益率数据，Pesando（1983）使用的是加拿大的债券数据。他们的研究都发现方差比VR大于1。这个结果得到了Singleton（1980）的支持，Singleton（1980）对（2）进行了规范统计检验（即他计算了VR合适的标准误差）。

如果我们假设EH中能增加时变期限溢价$T_t^{(n)}$，那么方差边界检验的违反可能是因为$T_t^{(n)}$具有可变性。如果是这种情形，那么期限溢价的可变性必须非常大才能使实证结果与方差边界检验结果完全相反。

然而，在上述早期方差边界检验中存在一些计量经济学问题。第一，如果利率序列具有随机趋势（即非平稳），那么方差无法定义，通常的检验统计量也就变得不是很合适。第二，即使利率序列是平稳序列，Falvin（1983）还是证明了如下结论：在通常使用的检验统计量中存在实质性的小样本偏差问题。为了克服后一个问题，Shiller（1989）（第13章）使用美国（1857—1988）和英国（1824—1987）的长期样本数据来对R_t和短期利率的移动平均（即完全预见长期利率R_t^*）进行比较。为了避免利率非平稳性所带来的问题，Shiller（1989）进行了如下回归：

$$R_t^*-R_{t-1}=\alpha+\beta\left(R_t-R_{t-1}\right)+v_t \tag{3}$$

Shiller发现，美国的$\hat{\beta}=1.156$（标准差＝0.253），英国的$\hat{\beta}=0.347$（标准差＝0.178）。就美国而言，我们能接受EH，因为$\hat{\beta}$与1之间的差异在统计上不显著；但对英国而言，在传统的显著性水平下，我们要拒绝EH。就整体而言，到期收益率长期数据集合的使用使得EH（具有常数期限溢价）在美国有较好的解释力，但在英国的解释里却较弱。

我们使用最新的数据画出了美国和英国R_t和R_t^*的图形，如图1a和图1b所示。从图形可以看出，在1979年之前，美国（图1a）的R_t和R_t^*在很长时间内大致平行。美国在1979年引入了货币基础控制政策一致持续到1982年，此后，通货膨胀率（从而利率）成下降趋势。EH在1979—1982年期间显然不成立，尽管在1982年后R_t的变化比R_t^*大。

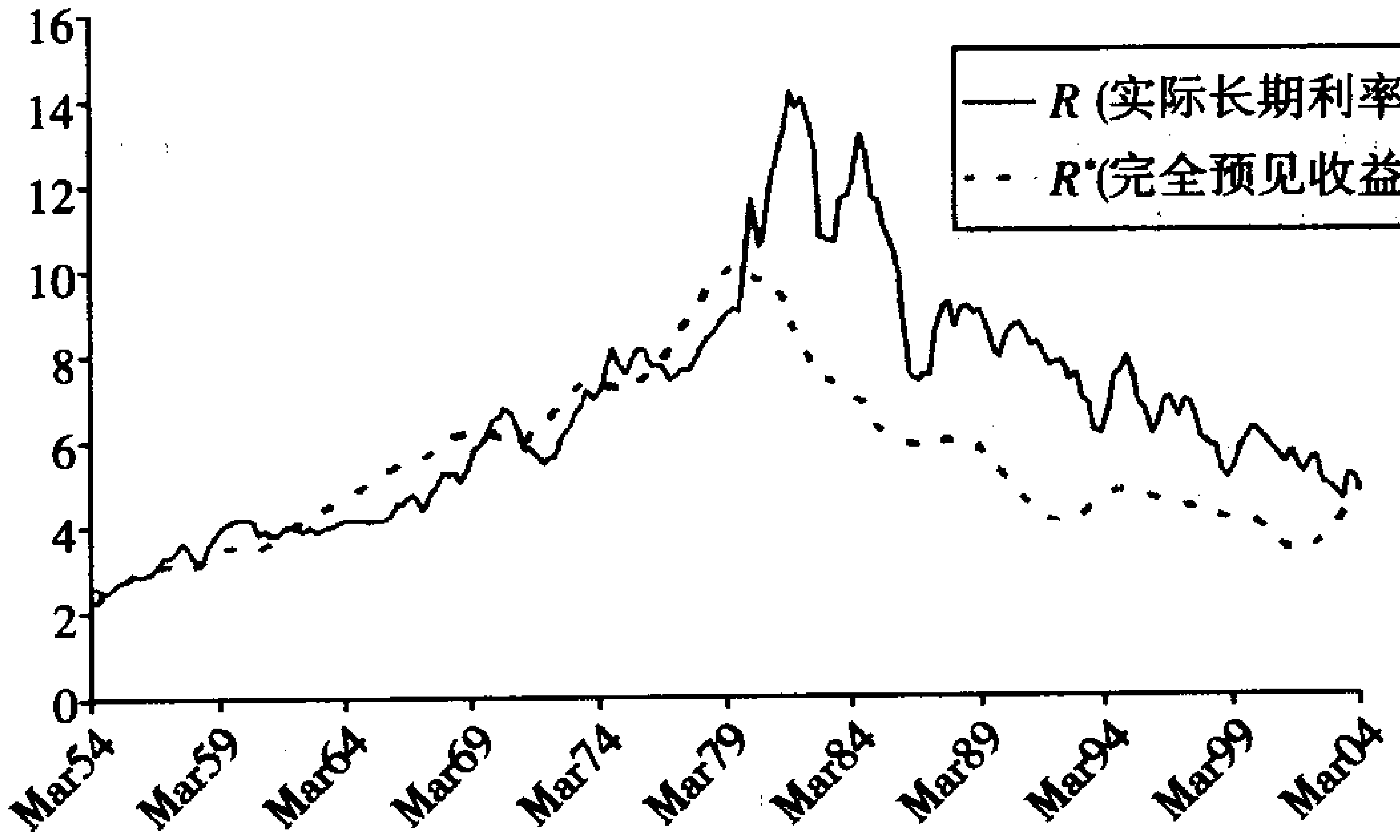
就英国而言（图1b），EH在1970—1985年期间很明显有非常大的问题，因为长期利率R_t

比完全预见长期利率 R_t^* 的波动幅度大得多，这种模式（尽管有时会出现中断的情形）一直持续到（历史上的）最后几年。在最后几年中，利率值较小且非常平稳。当然，长期利率的水平值非常接近于随机游走，两个图形使用的数据集合都比较短，因此我们的“因果关系实证”提供的仅仅是一个“棒球场式”（ball park）的估计而已。

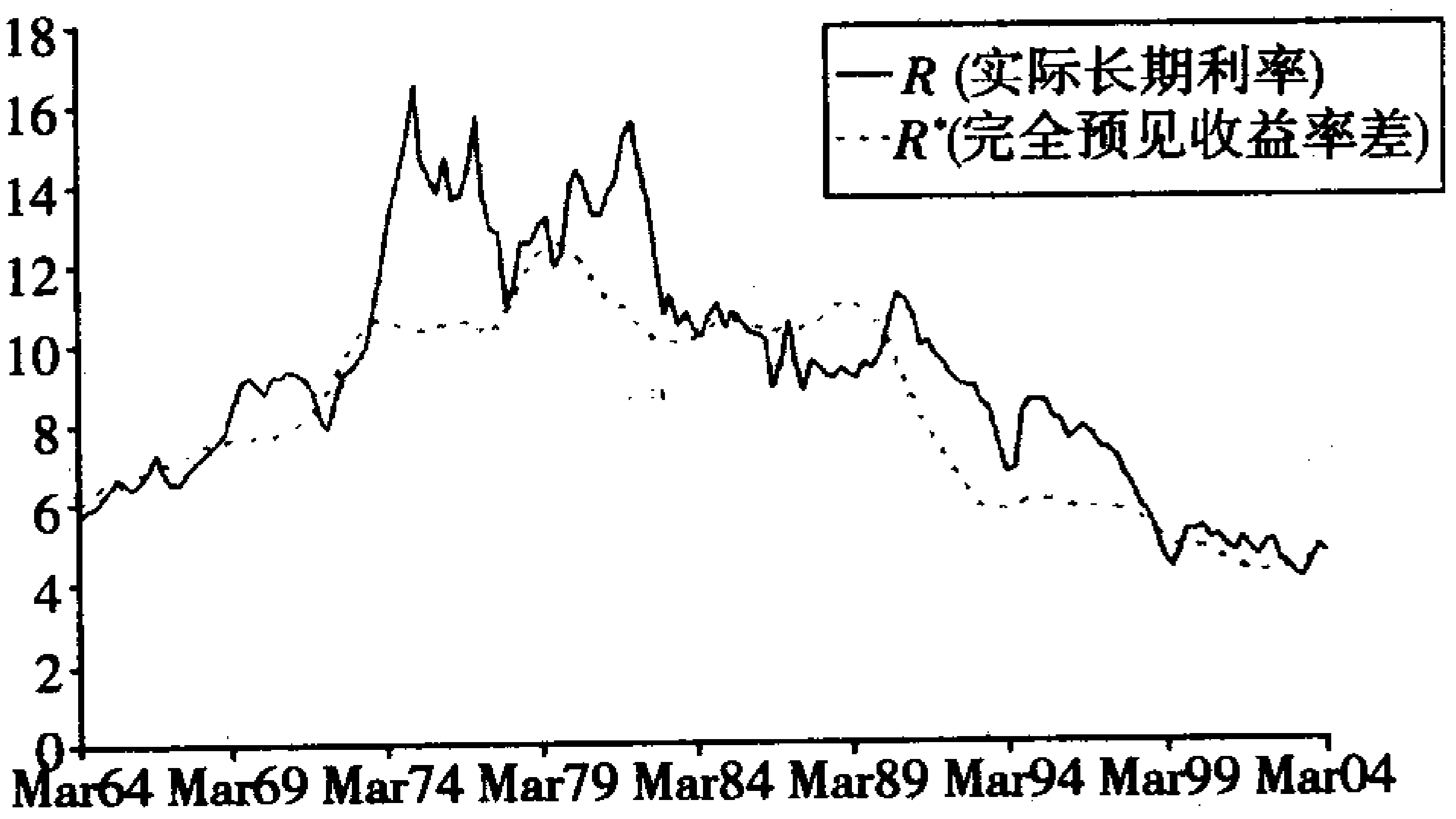
22.3 单方程检验

超额持有期收益率

根据 EH，超额持有期收益率（EHPR）应独立于 t 时刻的信息，任何流动性溢价都只依赖于期限，随着期限的增加而增加。如果我们对美国（1953—1997）1 到 5 年期债券的 HPR 取无条件平均值，那么（Cochrane（2000））$HPR^{(1)} = 5.8$（$\sigma_1 = 2.83$），$HPR^{(5)} = 6.4$（$\sigma_5 = 6.6$）。假定这个时期的平均利率为 5% 左右，超额 HPR 不会有很大的增加，但比 HPR 的标准差小得多。因此，债券的夏普比比股票的夏普比 $SR = EHPR/\sigma$（EHPR）小得多，不同期限债券的夏普比都为常数。从宽泛的角度来看，其他 G10 国家也有类似的结果。因此，初一看上去，EH 似乎是合理的。然而，对基于条件（而不是无条件）预期的检验进行考察后就会发现，EH 无法得到数据的支持。



(a)实际收益率和完全预见收益率(美国:1954—2004)



(b)实际收益率和完全预见收益率(英国:1964—2004)

图 1

远期利率回归

根据EH，t 时刻的单期远期利率（适用于从 n 到 $n+1$ 之间的时期）应是 $t+n$ 时刻单期到期收益率的无偏预测：

$$R^{(1)}_{t+n} = \alpha^* + \beta^* f^{(n)}_t + \eta_{t+n} \tag{4a}$$

$\beta^* = 1$。与方程（4a）一样对 $n=2, 3, \ldots$ 的不同远期利率进行回归就会发现 β^* 和 R^2 非常接近于1。然而不幸的是，这不能使EH有效，因为 r 和 f 都是非平稳序列，所进行的回归很有可能是伪回归。我们应该看到的一个更为合理且统计上更为有效的结果是：远期即期收益率能预测利率的变化，因为这些变量都是平稳的：

$$R^{(1)}_{t+n} - R^{(1)}_t = \alpha + \beta\ (f^{(n)}_t - R^{(1)}_t) + \eta_{t+n} \tag{4b}$$

$R^{(1)}_t = r_t$ 为当前的短期利率。像方程（4b）那样的回归分析给出的是非常不同的结果。对 $n=1$ 年期的债券而言，R^2 下降到0.005（1/2%）左右；对 $n=4$ 年期的债券而言，R^2 上升到0.75（标准差 = 0.30）。

方程（4b）的一个问题是，在期限较长时存在“数据损失”问题。例如，如果 $n=15$ 年，那么在样本的终端，我们就损失了180个月的观测值。方程（5）也存在序列相关性扰动。Backus、Foresi、Mozumdar和Wu（2001）建议使用方程（5）的一种替代方法。EH意味着

$$f^{(n)}_t = E_t R^{(1)}_{t+1} + rp^{(n)}_t \tag{5a}$$

$$f^{(n-1)}_{t+1} = E_{t+1} R^{(1)}_{t+n} + rp^{(n-1)}_{t+1} \tag{5b}$$

其中，$rp^{(n)}_t$ 为 $t+n$ 期的风险溢价，$f^{(n)}_t$ 为 n 期后的单期远期利率，它是 $t+n$ 期单期即期利率的无偏预测。注意，方程（5b）中的远期利率（从 $t+1$ 时刻开始，期限长度为 $n-1$）与 $R^{(1)}_{t+n}$ “在时间上是巧合的”。根据（5a）和（5b），运用迭代期望法则（即 $E_t E_{t+1} = E_t$），假设在远期利率期限溢价的基础上对EH进行回归检验，我们就得到：

$$f^{(n-1)}_{t+1} - R^{(1)}_t = \delta^{(n)}_0 + \delta^{(n)}_1\ (f^{(n)}_t - R^{(1)}_t) + \varepsilon^{(n)}_t \tag{6}$$

方程包含 $R^{(1)}_t$ 的目的在于使回归变量变成平稳序列，在EH下我们期望 $\delta^{(n)}_1 = 1$。方程（6）的优点在于：就10年的时期而言，被解释变量单期远期利率有119个观察值，不存在数据截断（即数据损失）问题。我们也避免了在方程（6）中使用重叠数据，而在方程（5a）中，重叠数据的使用倾向于“扭曲”小样本（GMM）估计的标准差。

Backus、Foresi、Mozumdar和Wu（2001）及Christiansen（2003）使用美国1976—1998年的数据发现，对 $n=1$ 到25年的期限而言，$\delta^{(n)}_1$ 约等于0.95，但与1之间存在显著差异。然而，当美国将1979—1982期间的货币基础控制政策取消后，以1987—1998年为样本期对 $n=1$ 到25年的期限而言，$\delta^{(n)}_1$ 大约为0.98～1.0，与1之间的差异不显著，因此这个结论为EH提供了支持。

因此，就期限超过一年的美国债券数据而言，一旦我们将1979—1982年货币基础控制试验期间利率波动较大的那些时期去掉后，EH在远期利率回归的基础上具有较强的解释力。

纯贴现债券/短期国债

早期对EH的大量实证研究使用的是3到6个月的短期国债数据（即纯（零息）贴现债券）。长期利率方程、短期利率方程和EHPR方程都非常简单。如果使用季度数据，那么我们有如下结论：

长期利率变化

$$(R^{(3)}_{t+1} - R^{(6)}_t) = \alpha_L + \beta_L S^{(6,3)}_t + \varepsilon_{t+1} \tag{7}$$

其中，$S^{(6,3)}_t = R^{(6)}_t - R^{(3)}_t$，$\alpha_L$ 是常数期限溢价。（注意，六个月的债券在三个月后就变成了三个月的债券。）

短期利率变化

$$\frac{1}{2}\Delta R_{t+1}^{(3)} = \alpha_S + \beta_S S_t^{(6,3)} + \varepsilon_{t+1} \tag{8}$$

它是完全预见收益率差 $S_t^{(6,3)*}$ = （1/2）$\Delta R_{t+1}^{(3)}$对实际收益率差 $S_t^{(6,3)}$ 进行回归的方程。

HPR 回归

$$H_{t+1} - R_t^{(3)} = c_0 + c_1\Omega_t + \varepsilon_{t+1} \tag{9}$$

因为以上三个方程右边的变量都是在 t 时刻到期，所以它们与 RE 预测误差 ε_{t+1}不相关，因此对这些回归方程使用 OLS 所得到是参数的一致估计值（但 GMM 要求对协方差矩阵进行纠正）。在 EH 下，我们会发现

$$H_0: \beta_L = \beta_S = 1 \quad 且 \quad c_1 = 0$$

使用上述方程中的任何一个进行回归都会得到相似的推断，因为估计参数之间存在精确的线性变换关系（如 $n = 2m$）。就美国 1963（1）—1983（4）的季度数据而言，Mankiw 和 Summers（1984）发现 $\beta_L = -0.407$（标准差 $=0.4$），但符号是错误的。Simon（1989）使用美国 1961—1988 年期间短期国债的每周数据发现 $\beta_S = 0.04$（标准差 $=0.43$），在 1972—1979 年期间的结果是 $\beta_S = 0.8$（标准差 $=0.34$），但它与 1 之间的差异在统计上不显著。两个研究都发现应该在较高的置信水平下拒绝预期假说。

Mankiw 和 Summers（1984）通过考虑如下可能性来寻找 EH 解释失灵的原因：市场参与者整体对 r_{t+1}的预期（由“精明”交易者的理性预期利率 E_t（r_{t+1}）和基于当前利率的简单而又天真的近视预测计划（即噪声交易者）加权平均而成）。如果 $\tilde{r}_{t+1}^{e}$表示市场的平均预期，那么 Mankiw 的假设为

$$\tilde{r}_{t+1}^{e} = wr_t + （1-w）（E_t r_{t+1}） \tag{10}$$

其中，$0 < \omega < 1$。因此 EH 就演变为

$$R_t = \frac{1}{2}r_t + \frac{1}{2}\tilde{r}_{t+1}^{e} + \phi \tag{11}$$

将方程（10）代入方程（11）并进行重新排列后就得到与方程（8）类似的方程，但$\beta_S = (1+\omega) / (1-\omega) > 1$。然而，嵌入这种“混合”预期计划并不需要借助于预期假说，因为 Mankiw 发现β_S 为负。在后一篇文章中，Mankikw 和 Miron（1986）对如下问题进行了考察：为什么在使用三个月和六个月的债券数据时，EH 在 1915 年之后表现得非常糟糕，但在 1890—1914 期间却表现得非常好。在 1915—1979 年期间的四个子样本上进行估计时，Mankiw - Miron 发现β_S 接近于零，R^2 也非常小（<0.06）。就 1890—1914 年（季度数据）而言，他们发现结果有所改进

$$\Delta r_{t+1} = \underset{(0.14)}{-0.57} + \underset{(0.18)}{1.51}\,(R_t - r_t) \tag{12}$$

$R^2 = 0.40$，尽管β_S 为正数，我们仍然不能解释$\beta_S = 2$ 的零假设。他们的建议是，在 1915 年之前的时期（在联邦储备系统建立之前），EH 绩效改善的原因在于 $E_t\Delta r_{t+1}$的可变性更强。在 1915 年之后，联邦储备银行试图使利率变得更加匀匀滑，这种行为可以表示为 $E_t\Delta r_{t+1} = 0$，即 r_{t+1}遵循随机游走（严格地讲是鞅）。因此，在 1915 年之后，收益率差对利率的未来变化没有任何预测功能。如果存在与收益率差相关的时变期限溢价，那么 OLS 得到的$\hat{\beta}_S$ 不是一致估计，即 $\text{plim}\hat{\beta}_S \neq \hat{\beta}_S$。这种“偏差”越小，$E_t\Delta r_{t+1}$的方差就越大。当 Δr_{t+1}无法预测时，σ^2（$E_t\Delta r_{t+1}$）$=0$，$\text{plim}\hat{\beta}_S = 0$。随着 $E_t\Delta r_{t+1}$的增加，$\hat{\beta}_S$ 的估计值接近于真实值。Mankiw - Miron 证明在 Δr_{t+1}的简单预测方程中，

$$\Delta r_{t+1} = \theta_1 \text{（}L\text{）} r_t + \theta_2 \text{（}L\text{）} R_t + \varepsilon_{t+1}$$

R^2 从 1890—1914 年期间的 0.4 下降到 1915 年后的零左右，这证实了上述猜测：EH 在 1915 年前表现得更好，因为利率的可预测性更强。

在期限较短的时期中，EH 表现糟糕的另外一个可能原因是，货币当局可能通过短期利率的增加来对长期短期收益率差的大小做出反应（后者给出的信号是未来预期通货膨胀率会更高，因此货币政策好似乎是矛盾的）。Kugler（2002）发现美国 1979—1982 年期间这个反应函数的参数发生了变化的证据，他将 EH 的拙劣表现归因于这种效应。

22.4　预期假说：案例研究

出于教学目的，我们首先详细考虑 Cuthbertson 和 Nitzsche（2003）对英国即期利率（最近由英格兰银行进行颁布，Anderson、Breedon 和 Deacon（1996），Anderson 和 Sleath（1999））的研究。即期利率通过“价格同时收集的”付息债券来进行计算，即期利率代表的是实际交易发生时的利率（除去非常小的经纪费用）。考虑的期限为 2，3，…，10，15，20，25 年。数据集合是从 1975 年 2 月开始的月度数据。另外，针对美国的假设检验是在 5% 的显著性水平下进行的。

当我们没有 $\ln R_{t+1}^{(n-1)}$ 的数据时，数据集合也可以使我们避免早期工作中所使用的长期利率逼近技巧，即用［$\ln P_{t+1}^{(n)} - \ln P_t^{(n)}$］来逼近 $h_{t+1}^{(n)} =$［$\ln P_{t+1}^{(n-1)} - \ln P_t^{(n)}$］。Bekaert、Hodrick 和 Marshall（1997）已经证明，这种逼近会导致 EH 回归检验产生偏差，对“长期利率变化的方程”而言，这种偏差尤其明显。对 EH 的所有研究都是在较长市场数据的末端进行的，所以我们的数据也存在一些逼近误差，因为收益率曲线是拟合的，尽管这样的操作肯定可以使原始数据中“异常值”的影响达到最小。

我们用一月利率和三月利率来表示短期利率，但是因为两种利率的定性结果并无差异，所以我们仅仅报告使用一月利率的结果。图 2 和图 3 是相对于一月短期国债利率而言的 2 年期利率和 25 年期利率，显然两个序列在较长的时期中一起运动，收益率差有相当大的变化。

核心方程是

$$(n-1)\ \{E_t R_{t+1}^{(n-1)} - R_t^{(n)}\} = \beta_L S_t^{(n)} + T_t^{(n)} \tag{13}$$

$$S_t^{*(n,1)} \equiv \sum_{i=1}^{n-1} \left(1 - \frac{k}{n}\right) E_t \Delta r_{t+i} = \beta_S S_t^{(n,1)} + \Phi_t^{(n)} \tag{14}$$

在常数期限溢价的 EH 假设下，$\beta_L = \beta_S = 1$。

单方程回归

根据长期利率方程回归结果（表 1）中利率差系数 β_L 的值，在 $n=2$ 到 10 年的期限上，我们不能拒绝 EH 的零假设（系数的值为 1），但当 $n=15$，20，25 时，我们要拒绝零假设。当 $n=3$ 到 6 时，系数 β_L 为正且接近于 1，接着开始下降，在更长的期限后变为负数。上述结果与 OLS 结果存在偏差是一致的，当 $n=15$，20，25 时，造成这一结果的原因是时变期限溢价。然而，这些结果更为惊人的一面是，β 有相对较大的标准误差，回归方程的 R^2 较小。除 $n=15$，20，25 外，我们也能接受零假设：利率差对长期利率在下一期没有任何预测能力。Bekaert、Hodrick 和 Marshall（1997）（表 6）的小样本结果表明，OLS 得到的 β_L 估计值是从均值开始往上偏斜的，当真实值 $\beta_L=1$ 时，经验分布值为 $\beta_L \approx 2$。这意味着，当 $n=15$，20，25 时，如果在我们的数据中广泛使用偏差纠正的方法，那么我们能在很高的置信水平下拒绝 EH。

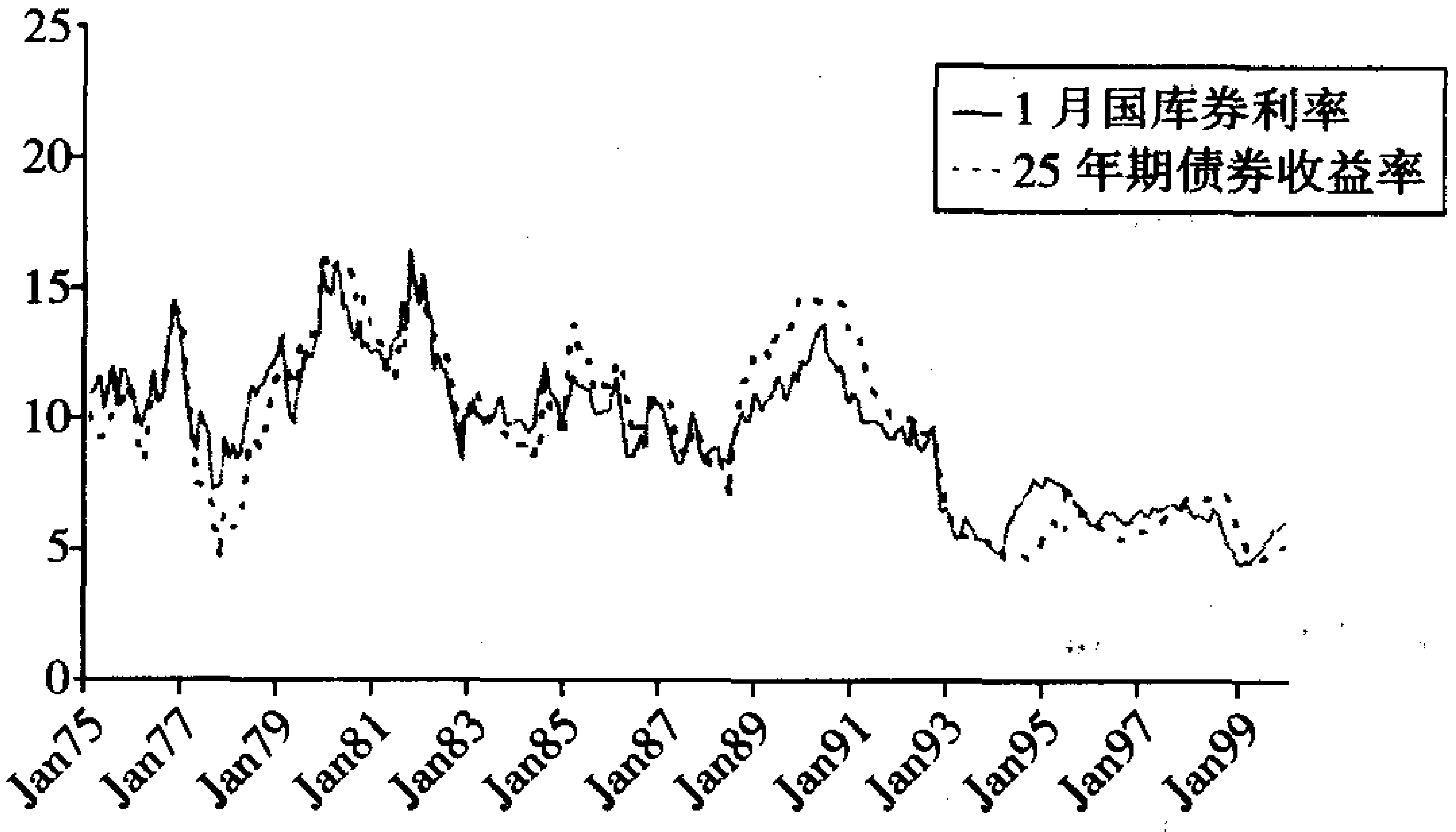


图 2　英国利率（2 年，1 月）

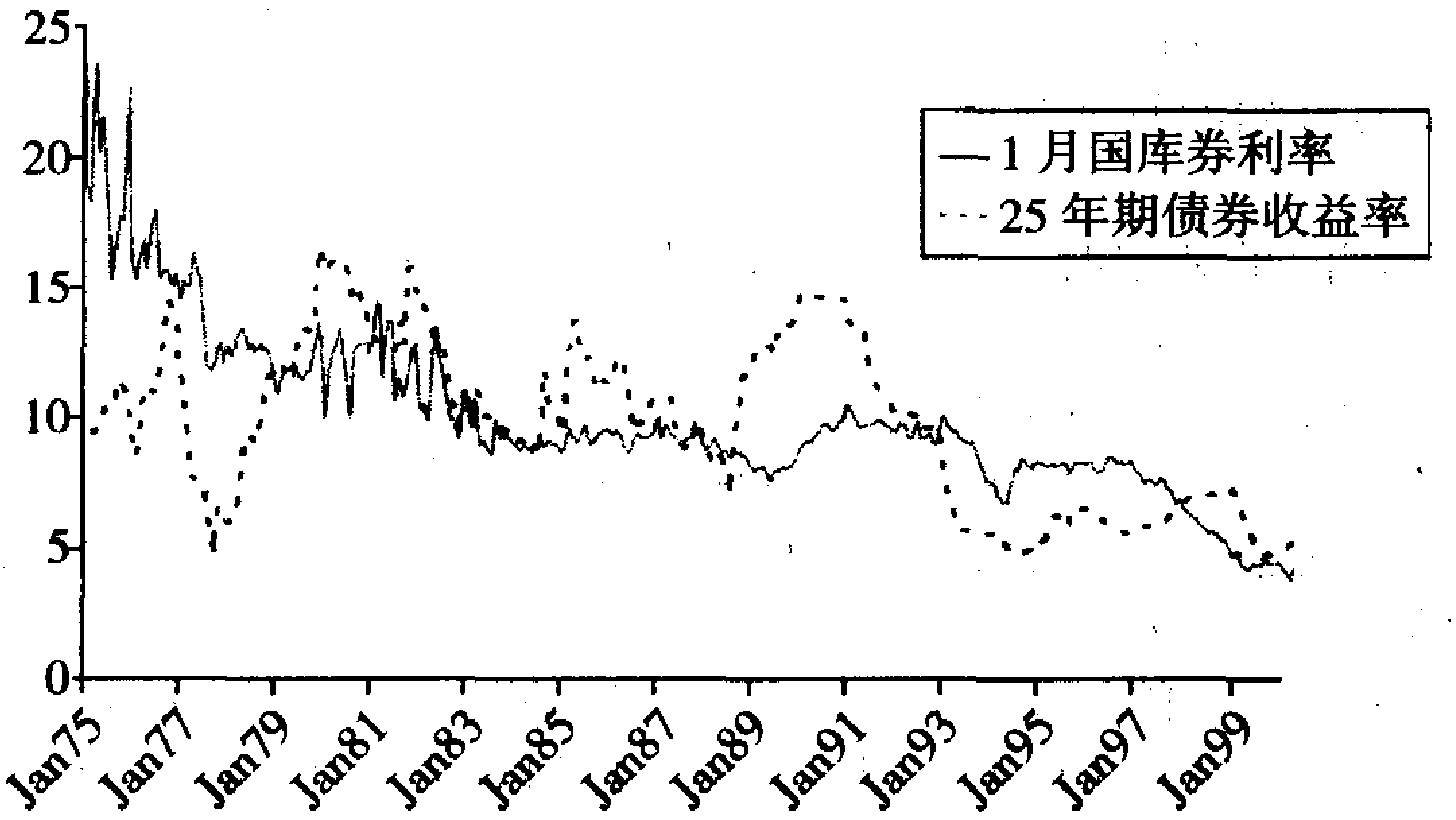


图 3　英国利率（25 年，1 月）

表 1　“长期利率方程”：$(n-1)\ R_{t+1}^{(n-1)} - R_t^{(n)} = \alpha + \beta_L\ (R_t^{(n)} - r_t)\ + \varepsilon_{t+1}$

利率期限	α		β		R^2
	系数	标准差	系数	标准差	
（2 年，1 月）	-0.5352	0.9189	0.4506	0.7136	0.0013
（3 年，1 月）	-0.9921	1.3083	0.9113	0.8619	0.0038
（4 年，1 月）	-1.4752	1.6742	1.1470	0.9934	0.0045
（5 年，1 月）	-1.9116	2.0365	1.1502	1.1172	0.0036
（6 年，1 月）	-2.2625	2.4053	0.9562	1.2374	0.0020
（7 年，1 月）	-2.5293	2.7808	0.6229	1.3540	0.0007
（8 年，1 月）	-2.7351	3.1601	0.2027	1.4660	0.0001
（9 年，1 月）	-2.9152	3.5407	-0.2666	1.5370	0.0001
（10 年，1 月）	-3.0958	3.9218	-0.7610	1.6751	0.0007
（15 年，1 月）	-4.4868	6.0238	-3.4832	2.1930	0.0085
（20 年，1 月）	-6.7220	9.6725	-7.5456	3.0920	0.0197
（25 年，1 月）	-9.7165	16.4440	-14.4373	4.6790	0.0312

注：方程是在 1976 年 1 月 ~ 1999 年 11 月期间用 GMM 方法进行估计的，因为存在异方差，所以我们用 White（1980）的方法对标准差进行了纠正。表中报告的是渐进标准差。

表2 "未来短期利率方程"：$\sum(1-i/n)\Delta r_{t+i}=\alpha+\beta_S(R_t^{(n)}-r_t)+\varepsilon_t$

样本期	长期利率期限（月）	α		β		R^2
		系数	标准差	系数	标准差	
1975年2月～1999年12月	$n=24$	-0.1559	0.3418	0.8794	0.1995	0.3032
1975年2月～1998年12月	$n=36$	-0.3722	0.4382	0.9791	0.2123	0.3749
1975年2月～1997年12月	$n=48$	-0.5959	0.4944	1.0746	0.2086	0.4676
1975年2月～1996年12月	$n=60$	-0.8110	0.5193	1.1158	0.1937	0.5442
1975年2月～1995年12月	$n=70$	-1.0175	0.4901	1.1251	0.1496	0.6097
1975年2月～1994年12月	$n=84$	-1.1999	0.4222	1.1391	0.1038	0.6698
1975年2月～1993年12月	$n=96$	-1.3703	0.3491	1.1816	0.0816	0.7360
1975年2月～1992年12月	$n=108$	-1.5519	0.3047	1.1533	0.0699	0.7810
1975年2月～1991年12月	$n=120$	-1.6416	0.2547	1.0978	0.0696	0.8105

注：方程是在第一栏中的样本期上用GMM方法进行估计的，因为存在异方差和n-1阶移动平均误差，所以对标准差要用Newey-West（1987）的方法进行修正。表只报告了$n_{\max}=120$（10年）所对应方程的估计结果，原因在于，为了与短期利率的未来变化相适应，必须对数据进行截断处理，因此我们的有限数据集合缺少自由度。

在未来短期利率方程（14）中使用完全预见收益率差所得到的结果（表2）为EH提供了更加有力的证据。事实上，$n=2\sim10$年的β_S系数几乎接近于1，（渐进）标准差小且合理。（注意，我们不能对期限超过10年的相应方程进行估计，因为存在数据损失的问题，正如我们在一个较长时期中对短期利率的未来变化进行建模时所遇到的情形一样。在下面的Var分析结果中，我们没有提及这个问题。）如果我们在Bekaert、Hodrick和Marshall（1997）中应用偏差纠正的方法，那么上述定性结果不会发生改变，因为使β_S向上偏斜的有限样本非常小（Bekaert等人也发现β_S的经验分布比渐进分布的分散程度要大一点点）。

Var分析

英国利率的单位根/协整性质已被广泛检验（例如，参见Taylor（1992），MacDonald和Speight（1991），Cuthebertson（1996），Cuthbertson、Hayes和Nitzsche（1996）），Var分析中$z=[S_t,\ \Delta r_t,\ h_t-r_{t-1}]$的每一个分量都是平稳序列。Var系数的$A$矩阵具有最大的特征值，（就我们的期限样本而言）特征值的取值范围为0.92～0.97，这意味着Var模型在预测未来短期利率时是动态平稳的。

Var模型的三个变量是$z_t=[S_t,\ \Delta r_t,\ h_t-r_{t-1}]$。超额持有期收益率$E_th_{t+1}-r_t$无时变性的限制条件，即$e3'A=0$只是在非常长的期限中（$n=15$，20，25年）在5%或更高的显著性水平下才被拒绝（表3）。这意味着，Var中的变量可能为这些较长期限的时变风险溢价提供了一种带有噪声的近似替代。

现在转向Var度量。使用Var后，理论收益率差S_t'是未来短期利率的预测值：$S_t'=f(A)z_t$。当$n=2$年时，实际收益率差S_t和理论收益率差S_t'的图像表明，两者的反应非常接近且合理（图4）；但当$n=25$年时，这种情形就不会发生（图5）。相对于未来短期利率的预期变化而言，收益率差具有明显的"过度反应"。

表4给出了实际收益率差S_t与理论收益率差S_t'之间关系的度量结果，结果表明，当$n=2\sim10$年时，我们无法拒绝Var的限制条件（使用Wald检验），实际收益率差与利率收益率差之间的相关系数非常接近1，标准差比率与1之间的差异在统计上不显著。从技术上讲，如果相关系数的渐进标准差是如此的小，那么这些相关系数应显著地不同于1。但是，在

Bekaert、Hodrick 和Marshall（1997）（表6）中，ρ（S_t'，S）的小样本经验标准差大约为0.10，这会导致我们的数据无法拒绝EH。

表3　　单期超额收益率具有时变特征吗？

利率期限	*Wald* 检验 H_0: $e3'A=0$ 统计量	 p－值
（2年，1月）	2.41	0.49
（3年，1月）	3.42	0.33
（4年，1月）	5.57	0.13
（5年，1月）	6.69	0.08
（6年，1月）	5.89	0.12
（7年，1月）	4.30	0.23
（8年，1月）	3.03	0.39
（9年，1月）	2.39	0.50
（10年，1月）	2.31	0.51
（15年，1月）	7.83	0.05
（20年，1月）	8.37	0.04
（25年，1月）	8.06	0.04

注：样本期为1976年1月~1999年11月。标准差是异方差稳健的（White（1980））。无时变（单期）期限溢价的零假设为：H_0: $e3'A=0$。检验统计量渐进服从自由度为3的χ^2分布（5%和10%显著性水平的临界值分别为7.815和6.25）。

当$n=2\sim10$时，我们的结果与如下结果形成鲜明对比：美国的数据表明，在较大期限范围的末端，可以很“干脆”地拒绝EH（Campell和Shiller（1991），Bekaert、Hodrick和Marshall（1997））。然而，当$n=15$，20，25时，英国数据的Var计量也表明应拒绝EH，这与相应期限的时变溢价的存在性一致（即拒绝限制条件$e3'A=0$）。

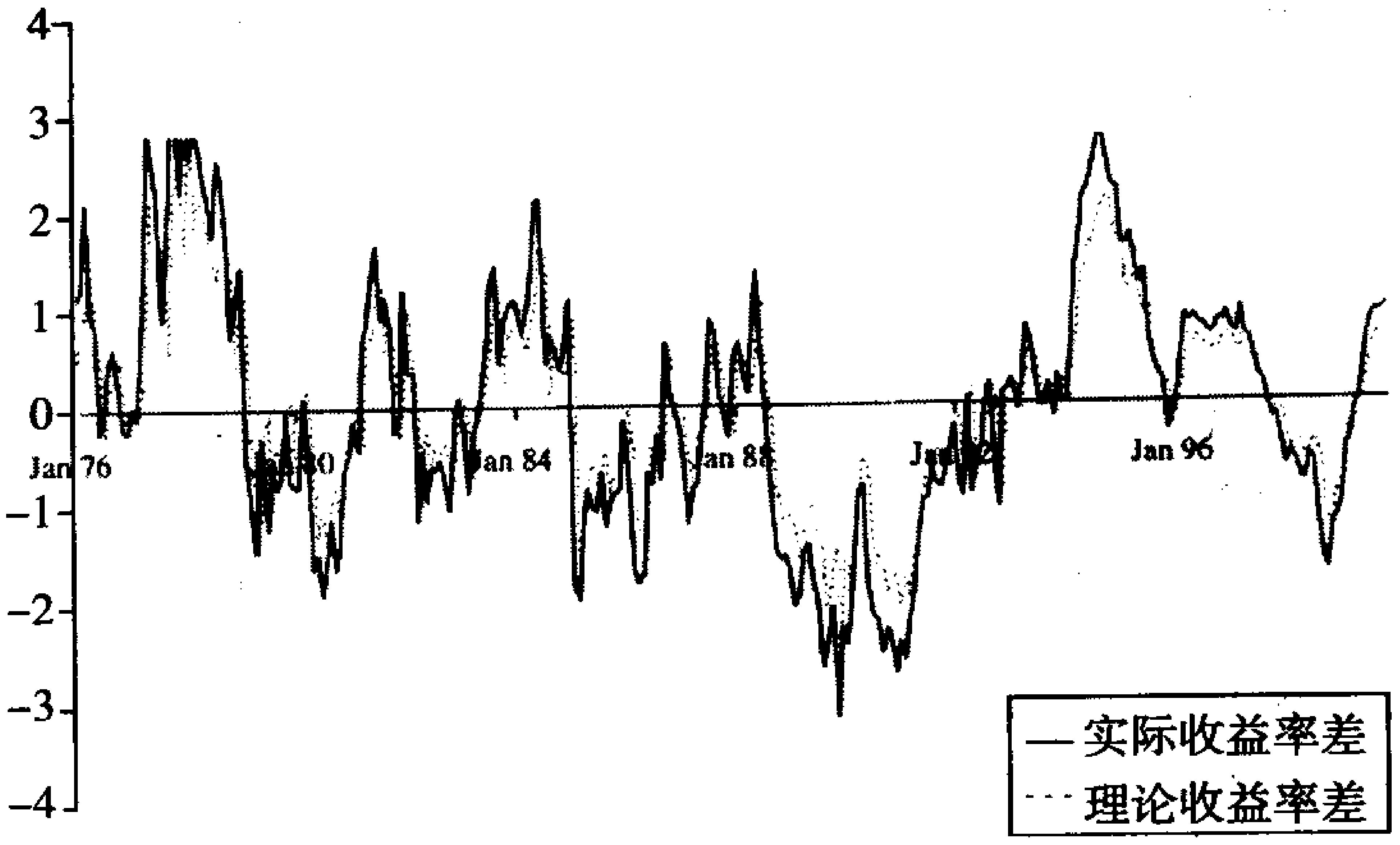


图4　英国的实际收益率差和理论收益率差（2年，1月）

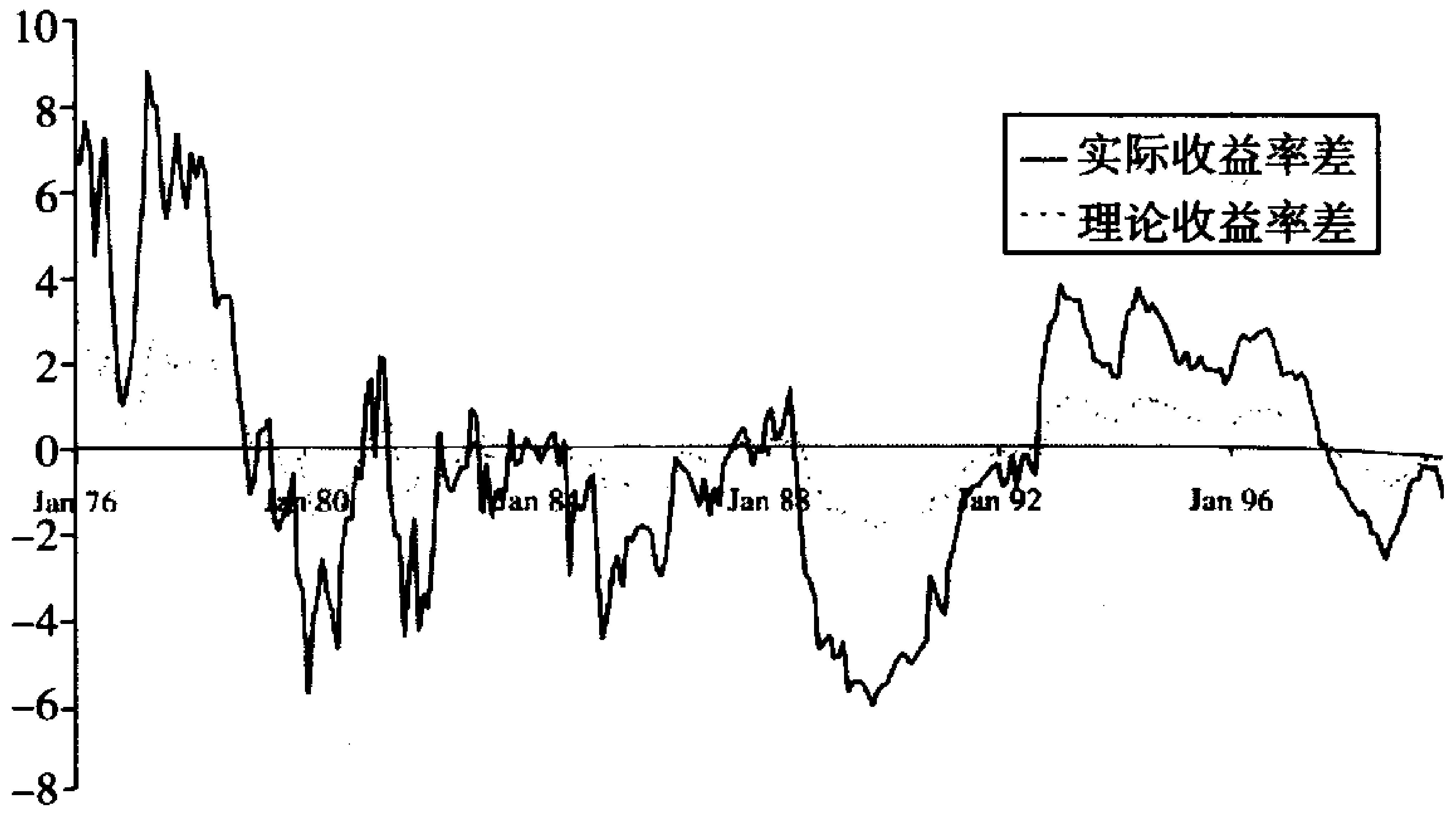


图5　英国的实际收益率差和理论收益率差（25年，1月）

表4　实际收益率差 S_t 和理论收益率差 S'_t

利率期限	Wald检验 $H_0: S_t = S'_t$		$\rho(S_t, S'_t)$		$\sigma(S'_t)/\sigma(S_t)$	
	值	p-值	值	p-值	值	p-值
（2年，1月）	2.43	0.49	0.9043	0.0092	0.7527	0.2116
（3年，1月）	3.25	0.35	0.9276	0.0074	0.8059	0.2321
（4年，1月）	4.88	0.18	0.9385	0.0068	0.8236	0.2619
（5年，1月）	5.67	0.13	0.9442	0.0063	0.8213	0.2666
（6年，1月）	5.04	0.17	0.9474	0.0058	0.8099	0.2772
（7年，1月）	3.80	0.28	0.9494	0.0053	0.7936	0.2862
（8年，1月）	2.76	0.43	0.9508	0.0050	0.7736	0.2496
（9年，1月）	2.22	0.53	0.9508	0.0048	0.7496	0.3015
（10年，1月）	2.22	0.53	0.9532	0.0047	0.7242	0.3062
（15年，1月）	8.20	0.04	0.9595	0.0049	0.5865	0.2885
（20年，1月）	10.63	0.01	0.9608	0.0054	0.4445	0.2043
（25年，1月）	22.55	0.00	0.9507	0.0094	0.3268	0.2015

注：样本期为1976年1月~1999年11月。零假设：实际收益率差 S_t 是短期利率未来变化（即理论收益率差 S'_t）的充分统计量，记为 $H_0: S_t = S'_t$。零假设意味着，用Wald统计量进行检验的Var的矩阵存在交叉方程限制。Var中的A变量为 $Z = [S_t, \Delta r_t, h_t - r_{t-1}]$。检验统计量渐进服从自由度为3的 χ^2 分布（5%和10%显著性水平对应的临界值分别为7.815和6.25）。

表 5　未来短期利率（er_{t+1}）方面的“消息”和单期收益率（eh_{t+1}）

利率期限	ρ（er_{t+1}，eh_{t+1}）		σ（er_{t+1}）/σ（eh_{t+1}）	
	统计量	标准差	统计量	标准差
（2 年，1 月）	-0.9931	0.0186	0.9512	0.0742
（3 年，1 月）	-0.9982	0.0101	1.0072	0.0797
（4 年，1 月）	-0.9994	0.0086	1.0509	0.0808
（5 年，1 月）	-0.9993	0.0083	1.0704	0.0795
（6 年，1 月）	-0.9979	0.0166	1.0656	0.0777
（7 年，1 月）	-0.9941	0.0314	1.0444	0.0765
（8 年，1 月）	-0.9870	0.0521	1.0164	0.0772
（9 年，1 月）	-0.9759	0.0773	0.9892	0.0827
（10 年，1 月）	-0.9607	0.1049	0.9673	0.0952
（15 年，1 月）	-0.8300	0.2070	0.9680	0.2051
（20 年，1 月）	-0.6492	0.2292	0.9831	0.2114
（25 年，1 月）	-0.4791	0.2288	0.8128	0.1416

注：样本期为 1976 年 1 月 ~1999 年 11 月。零假设：预期单期超额收益 $eh=h_{t+1}-E_th_{t+1}$ 的唯一来源是未来短期利率（er）方面的消息。零假设意味着：ρ（er，eh）$=1$ 且 σ（er）/σ（eh）$=1$。Var 中的变量为 $Z=[S_t,\ \Delta r_t,\ h_t-r_{t-1}]$。

消息与期限溢价

限制条件 $e3'A=0$（$n=15, 20, 25$）（$n=15, 20, 25$ 年，表 3）的拒绝意味着，单期风险溢价 $T_t^{(n)}$ 是时变的。然而，在表 4 中，我们将实际收益率差 S_t 与理论收益率差 S_t' 进行了比较，结果表明，平均未来溢价 $\Phi_t^{(n)}=(1/n)\sum T_{t+i}^{(n)}$ 的变化相当小。我们能将这两种证据协调起来吗？

在表 5 中，我们将预料不到收益率 $eh_{t+1}\equiv h_{t+1}-E_th_{t+1}$ 与利率 er_{t+1} 未来变化方面的“消息”进行了比较。对小于 10 年的所有期限而言，标准差比率 σ（er）/σ（eh）和相关系数 ρ（er，eh）分别接近于 $+1$ 和 -1，这意味着预期不到收益率 eh_{t+1} 的绝大部分变化都是由未来短期利率 er_{t+1} 方面的“消息”所导致的，只有非常小的部分是由未来平均期限溢价 $\Phi_t^{(n)}$ 方面的“消息”所导致的。对 15 年、20 年、25 年的期限而言，相关系数有很弱的递减趋势，但除了 25 年期限外，我们仍然不能拒绝相关系数为 -1 的假设，而它的标准差比率为 0.8128（渐进标准差为 0.14）。

如果 $eT_{t+1}^{(n)}\approx 0$，那么 Var 中超额收益率方程（即第三个方程）的“$1-R^2$”给出的是超额持有期收益率中由未来短期利率方面消息的变化所导致的部分。Var 中方程的“R^2”（表 6）的变化范围为 0.01 ~0.05（除 $n=20, 25$ 之外，两个期限所对应的值分别为 0.11 和 0.14）。这意味着未来短期利率变化方面的消息能解释事后超额收益率几乎所有（即大约 95% ~99%）的变化，预期超额收益率只有非常小的变化。当然，后者要满足一个重要条件：Var 中变量可以充分模拟预期收益率的变化。表 6 也给出了收益率差、短期利率变化和超额持有期收益率（Var 方程）的 Ljung - Box 统计量结果，对 20 年和 25 年期限而言，只有最后一个方程才存在序列相关问题（这意味着期限溢价存在一些变化，而用已经包含的变量无法对这些变化进行很好的建模）。

单期期限溢价 $T_t^{(n)}$ 的变化对（$n=10, 15, 25$ 的）单期收益率的效应在统计上确实显著，因为我们无法拒绝 $e3'A=0$（表 3），因此这些效应导致了相应期限上 β_L 的有偏估计和 EH 被拒

绝的结果。然而，在长期限中短期利率的未来变化（表2）依赖于 $T_{t+1}^{(n)}$（$i=1，2，\cdots$）所有预期值的平均，在 $T_{t+i}^{(n)}$ 的现值公式中 $T_{t+1}^{(n)}$ 仅有 $1/n$ 的权重。我们的结果表明，$T_{t+i}^{(n)}$ 并没有强的持久性，因此，相对于未来短期利率的预期变化而言，平均未来风险溢价 $\Phi_t^{(n)}=(1/n)E_t\sum T_{t+i}^{(n)}$ 很小。因此，在多期中，未来短期利率的变化主导了未来平均风险溢价 $\Phi^{(n)}$ 的变化，即使期限很长，这一结论也同样成立。这也是如下现象的原因：对 S_t 和 S_T' 进行比较的结果（表4）表明，在时变风险溢价的平均值可以忽略的零假设下，EH 得到了广泛的支持。

表6　　Var 的描述统计量：$Z=[S_t，\Delta r_t，h_t-r_{t-1}]$

利率	Q－统计量						R^2		
（20年,1月）	S_t-eqn		Δr_t-eqn		$(h_t-r_{t-1})-eqn$		S_t	Δr_t	(h_t-r_{t-1})
	Q(1)	Q(6)	Q(1)	Q(6)	Q(1)	Q(6)			
（2年,1月）	0.49	4.38	0.63	10.3	0.02	8.09	0.8378	0.1127	0.0097
（3年,1月）	0.34	3.93	0.73	10.5	0.07	10.4	0.8758	0.1160	0.0165
（4年,1月）	0.34	4.10	0.79	10.9	0.12	10.2	0.8940	0.1121	0.0274
（5年,1月）	0.40	3.89	0.82	11.3	0.13	8.08	0.9036	0.1064	0.0321
（6年,1月）	0.50	3.89	0.82	11.5	0.11	6.33	0.9087	0.1003	0.0283
（7年,1月）	0.65	4.22	0.80	11.6	0.08	5.48	0.9117	0.0948	0.0200
（8年,1月）	0.82	4.57	0.78	11.6	0.06	5.14	0.9139	0.0900	0.0121
（9年,1月）	0,97	4.64	0.74	11.5	0.04	5.00	0.9161	0.0860	0.0076
（10年,1月）	1.08	4.32	0.71	11.4	0.02	5.00	0.9186	0.0825	0.0075
（15年,1月）	0.98	1.84	0.43	9.59	0.81	10.6	0.9318	0.0631	0.0547
（20年,1月）	0.12	1.54	0.12	6.90	3.72	20.2	0.9332	0.0344	0.1118
（25年,1月）	0.08	6.68	0.05	5.99	2.14	35.6	0.9163	0.0251	0.1413

注：样本期为1976年1月～1976年11月。对所有期限而言，Var 的滞后长度选择为 $p=1$。报告的 Ljung－Box 统计量的滞后长度为1个月和6个月，临界值（5%的置信水平）分别为3.84和12.59。

小结：英国数据

对“短期”债券（$n=2\sim10$ 年）而言，英国数据支持具有常数期限溢价的 EH，因此，对这些期限而言，持有期收益率无法预料的变化只依赖于对未来短期利率预测值的修正。

当考虑时变期限溢价时，对 EH 的检验要用三个变量的 Var，这三个变量不但包括收益率差和短期利率的变化两个变量（正如早期工作中所使用的那样），还包括超额持有期收益率变量，后一个变量捕捉了（平稳）期限溢价的运动特征。通过这些修正（正如 Tzavalis 和 Wickens（1997）所建议的那样），我们的结果支持（平稳）时变期限溢价存在的结论，但只是对长期限债券（$n=15，20，25$ 年）而言这个结论才成立，该时变期限溢价对单期超额持有期收益率存在影响。

然而，单期期限溢价 $T_t^{(n)}$ 不具有持久性，因此，相对于未来短期利率预期值的变化而言，单期时变期限溢价对平均期限溢价 $\Phi_t^{(n)}=(1/n)E_t\sum T_{t+i}^{(n)}$ 的影响较小。因此，这个时变期限溢价不会导致 β_S 的 OLS 估计值和 Var 模型存在很大的偏差。后面一个结果也是一致的，因为收益率差是未来短期利率的最优预测值，与 EH 所建议的一样（即具有近似时间不变性的平均期限溢价）。

然而，当 $N=10，15，20$ 时，时变期限溢价的存在性足以导致长期利率变化方程所表示的 EH 遭到拒绝（表2），因为单期期限溢价 $T_t^{(n)}$ 对长期利率的变化存在直接影响，收益率差方程中 β_L 的 OLS 回归系数可能存在非常严重的偏差。因此，通过在非常长的期限中引入时变期限溢价，

我们就能将对 EH 用英国数据进行的各种检验结果协调起来。

22.5 以前的研究

在期限谱的短端，Cuthbertson（1996）使用双变量 Var$[z=(S_t,\Delta r_t)]$和英国的即期利率数据对一年期上的 EH 进行了考察，而 Cuthbertson、Hayes 和 Nitzsche（1998，2000）使用的是德国数据，Cuthbertson 和 Bredin（2000，2001）使用的是芬兰的数据，他们都发现了支持 EH 的证据。（这些结果唯一的不足之处是交叉方程的 Wald 检验偶尔也会出现失灵的情形，参见下面要叙述的可能偏差。）

Taylor（1992）使用的是英国 1985 年 1 月 ~ 1989 年 12 月期间相对较长期限（5 年、10 年和 15 年）的数据。Taylor 所报告的结果是强烈地拒绝 *Wald* 限制（p 值为 0.00），因为理论方差比应等于 1，而实际的最小方差比为 1.5（*s. e.* =0.14）。这些结果与 Cuthbertson 和 Nitzsche（2003）针对 2 ~ 10 年期所报告的结果形成鲜明对比。结果存在差异的原因是，Taylor 使用的是到期收益率而不是即期收益率。

美国和其他国家

对美国而言，大量证据都支持 EH。Campell 和 Shiller（1991）及 Bekaert、Hodrick 和 Marshall（1997）对月度即期利率数据（对付息债券使用三次样条函数来进行拟合）应用 Var 方法的文献进行了综述。一般而言，对非常宽泛的期限（1 ~ 12 个月、2、3、4、…、10 年）而言，他们所发现的证据不支持 EH。尽管收益率差在正确的方向上能对短期利率的未来变化进行预测，收益率差的实际运动还是比“EH 是正确模型”的零假设所需要的运动大得多。这就是过度反应假说，它意味着实际收益率差不是短期利率未来变化的无偏预测。例如，Campell 和 Shiller（1991）使用的是美国 1946—1987 年期间 5 年及 5 年以下政府债券的月度数据，期限包括 1、2、3、4、6、9、12、14、36、48、60 个月。因此，他们使用的数据位于期限谱的短端。宽泛地说，对期限小于 1 年的情形而言，在完全预见收益率差 S_t^* 对实际收益率差 S_t 的回归分析中，他们几乎没有找到支持 EH 的证据，β 值位于 0 ~ 0.5 的区域而不是接近于 1。类似地，对期限小于 1 年的情形而言，Corr(S_t, S_t')相对较小，取值范围为 0 ~ 0.7，$\mathrm{VR}=\sigma^2(S_t)/\sigma^2(S_t')$的取值范围为 2 ~ 10。对期限为 4 年和 5 年的情形而言，Campell 和 Shiller（1991）发现了一些支持 EH 的证据，因为方差比率 VR 及 S_t 和 S_t' 之间的相关系数接近于 1。然而，Campell 和 Shiller 没有对 Var 交叉方程限制直接进行检验，但 Shea（1992）却对此进行了检验，一般而言；Shea 所发现的结果是拒绝 EH。

Bekaert、Hodrick 和 Marshall（1997）对 Campell 和 Shiller（1991）重新进行了研究，美国的数据表明，在长期利率方程中收益率差系数的估计值为 −0.8（1 年期）到 −2.3（5 年期），在未来短期利率方程中的估计值为 0.33（1 年期）到 0.57（5 年期）。根据渐近 p 值，EH（对 $n=1,2,5$ 年而言）在 4% 的显著性水平遭到拒绝。Var 方法所得到的结果是，S' 和 S 之间的相关系数为 1 的零假设不会遭到拒绝，S' 和 S 的标准差比率为 1 的零假设明显遭到拒绝（对渐近检验而言，显著性水平为 3%），对三种期限而言，标准差比率大约为 0.4。

Bekaert、Hodrick 和 Marshall（1997）还对上述检验在 EH 为真的零假设下的小样本性质进行了蒙特卡罗模拟研究（并对短期利率使用 VAR - GARCH 数据生成过程）。他们发现，在长期利率回归方程中，收益率差的回归系数严重向上偏斜（对 5 年期情形而言，当真实的 β_L 为 1 时，OLS 的估计结果为 +2.6），在未来短期利率回归方程中，收益率差的回归系数向上偏斜的程度较轻（对 5 年期的情形而言，β_S 的估计值为 1.4，而真实值为 1）。尽管小样本经验分布的分散程度比渐近分布高，在“长期利率”方程和“短期利率”方程的基础上，上述小样本结果增强了对 EH 的拒绝。他们所发现的一个有趣结果是，在 Var 方法中 $\rho(S_t', S_t)=1$ 和 $\mathrm{SDR}=\sigma(S_t')/\sigma(S_t)=1$ 的估计结果几乎没有偏差（与渐近值相比，尽管 SDR 的经验标准差有点偏小）。然而，他们再一次

发现，这些小样本蒙特卡罗模拟结果增强了美国数据对 EH 的拒绝程度。

在后续研究中，Bekaert、Hodrick 和 Marshall(2001)使用蒙特卡罗模拟方法在“EH 成立”的零假设和短期利率数据生成过程存在区制转移的假设下来考察“长期利率变化”方程和“未来短期利率变化”方程中收益率差系数的小样本性质。模型对“回归系数存在偏差”的“比索问题”(Peso problems)进行了模拟。他们发现，比索效应导致回归系数向上出现较大的偏差，并不令人感到奇怪的是，比索效应也增加了小样本分布的分散程度(与渐近分布进行比较)。但是，拒绝 EH 的证据变得非常无力。当他们在模型中增加规模较小的时变期限溢价(正确定价)时，在数据的解释方面，β_S 和 β_L 的小样本分布分散程度更大、偏斜更加严重、偏差也更大(例如，在长期利率方程中，β_L 的一些负值出现在 2.5% 的尾部中)。然而，蒙特卡罗模拟结果不能解释 5 年期或更长期限利率的回归结果。

Harris(2001)使用美国的面板数据来考察时变期限溢价导致长期收益率方程(13)中收益率差系数 β_L 的 OLS 估计值产生偏差的可能性。单因子期限溢价模型的使用使 OLS 系数估计变成无偏估计，期限溢价由面板数据回归中的时间固定效应来捕获。如果不考虑时变期限溢价，那么(13)中收益率差系数的平均值为 $\beta_L = -3.2$(1～13 年期限上的平均值)，而面板数据中的平均值为 -1.98(*s. e.* =0.213)。这意味着时变期限溢价的出现导致(13)中 OLS 估计值出现了偏差，考虑时变期限溢价后仍然要拒绝 EH，因为 β_L 的系数与 1 在统计上仍然存在显著差异。

Mankiw 和 Miron(1986)认为，从实证角度来看，EH 在货币目标政策情形中比匀滑利率情形中的表现好。Kugler(1988)使用的是美国、德国和瑞士的 1 月和 3 月欧元存款利率的月度数据，结果发现，只是在德国数据(1974.3～1986.8)中 EH 才获得支持，这基本上与 Mankiw－Miron 假说一致。类似地，Engsted(1996)使用的是丹麦货币市场数据和更长期限的债券数据(Engsted 和 Tanggaard 1995)，结果发现，如果相对而言的利率变化幅度较大(即 1992 年 ERM“危机时期”之后的时期)，那么支持 EH 的证据就变得更加确凿。Mankiw 和 Miron(1986)分析的预期结果是：如果利率稳定导致短期利率出现随机游走行为，那么短期利率的预期变化为零，收益率差对未来短期利率没有任何预示功能，这与 EH 形成对比(也可参见 Rudebusch(1995))。

由 Mankiw 和 Miron(1986)可以看出，对 EH 的计量检验要求短期利率预期变化的可变性足够强，尽管如此，如下情形还是可能发生：因为时变期限溢价存在，所以非常大的(预料不到的)变化可以增加代理人对持有期风险的认知，从而使 EH 变得无效(参见 Engle、Lilien 和 Robins(1987)，Hall、Anderson 和 Granger(1992)，Tzavalis 和 Wickens(1995))。时变期限溢价(用三变量 Var 中的持有期收益率来对其进行建模)的出现似乎有助于解释美国数据中的结果，在英国数据中长期期限溢价的效应似乎“更强一些”(参见 Tzavalis 和 Wickens(1997))。

与期限谱长端的结果相对比的是，Tzavalis 和 Wickens(1997)使用的是美国(1970—1986)3、6、12 月短期国债的月度数据，使用 Var 方法所得到的结果强烈支持 EH(只要将 1970—1982 的货币基础控制期忽略掉)，这与 Longstaff(2000a)使用美国重购协议利率(从隔夜到三个月的利率)所发现的结果类似。在用美国数据对 EH 的有效性进行检验时，对 1979—1982 时期中的利率进行建模就变得非常重要，因为这个时期中的利率具有高度波动性，所以人们预期时变期限溢价存在。Driffell 和 Sola(1994)使用了 Var 框架(只使用美国 3 个月利率和 6 个月利率)和 Hamilton(1988)两种状态的 Markov 区制转换模型。他们在未来短期利率方程 $S_t = (1/2)E_t\Delta r_{t+1} + u_t$ 中也引入了白噪声误差项，这就使得 Var 交叉方程的约束性更少(与 $u_t = 0$ 的情形相比较)。在发现“*EH* 在 1979—1982 期间不成立”后，他们重复使用 Var 程序，假设系统在 1979—1982 期间处于不同状态。现在，EH 的交叉方程限制变得可接受了，Hamilton 程序给出的“系统在 1979—1982 期间处于不同的状态”的概率为 1。因此，Hamilton 滤波几乎等价于在方程中为这个时期引入了如下虚拟变量：

$$\{\cdots,0,0,1,1,1,\cdots,0,0,\cdots\}$$

Tzavalis 和 Wickens(1995)沿着稍微不同的路线对 1979—1982 进行了考察，结果表明，GARCH 模型能充分捕获超额持有期收益率在这一时期上的条件波动率。然而，如果在

GARCH(1,1)中加入一个虚拟变量(虚拟变量在1979—1982的时期上取值为1),那么条件波动率不再对$(h_{t+1}-r_t)$产生“均值”效应,尽管在这个方程的残差中GARCH效应仍然很显著(但持续性更弱)。因此,使用美国数据对EH的任何检验都要谨慎地对待1979—1982时期的影响。在利率高波动时期中EH倾向于“失灵”的现象在其他国家和地区也会出现,如香港(1992—2000)的短期利率(如3~12个月的期限)。Gerlach(2003)发现,在短期利率变化方程中增加GARCH协方差项后,该项在统计上显著,这意味着香港市场存在时变期限溢价现象,检验结果不会拒绝$\beta_S=1$的零假设(尽管β_S的度量存在很大的误差,标准差也很大)。

为什么结果存在分歧呢?

我们能对不同期限和不同国家的EH结果存在分歧的现象进行解释吗?一般而言,与Var方法相比(特别是与Var交叉方程限制遭到拒绝的结果相比),完全预见回归方程所给出的结果是强烈支持EH。一个原因是,使用(14)的完全预见回归方程间接考虑了潜在的未来事件(代理人知道这些事件,但计量经济学家却不知道)对预期的影响,而Var方法要求代理人和计量经济学家在t时刻或之前就知道信息集合。政府当局的货币政策对短期金融工具市场经常产生重大影响,中央银行的利率变化政策对择时交易的“第二次猜测”有着重要影响。在政府干预(影响)时期中,纯粹的“后瞻式”回归可以视为短期利率未来变化的糟糕预测,然而,理性预期假设$r_{t+1}=E_t r_{t+j}+\omega_{t+j}$只需要无偏性,有偏性所导致的效应很小。因此,出于这个原因,人们预期完全预见回归的表现比Var方法好,对EH的支持程度也更强一些(如果EH事实上为真)。

数据收集的频率、利率报价同步的程度(即利率是否在一天的同一时刻进行报价?)、计算债券收益率差或价格变化时所使用的逼近技巧都可以解释上述研究中的矛盾结果。最为重要的是,Bekaert、Hodrick和Marshall(1997)表明,检验统计量的小样本偏差可以解释Wald检验和长期利率方程中的不同推断结果,但Var方法中的方差比和相关系数统计量却不行。因此,“检验统计量”的蒙特卡罗模拟方法和自助法对实证分歧结果的解释有一定的帮助。

Kozicki和Tinsley(2001)增加了“短期利率的标准单变量时间序列表示(与随机游走表示一样)是有偏的”可能性,因为他们考虑了通货膨胀政策目标的市场认知转换,这对未来短期利率的长期预测给出了一个边界。因此,他们对长期限中未来短期利率的预测施加了“终点”限制,代理人缓慢地学习新的货币政策。通过这些预测再加上EH,他们就能对每一个时点上的长期利率进行预测。就美国1960—1995年期间的数据而言,预测的长期利率和实际长期利率非常接近(见图2~4)。短期利率的VECM预测有一个误差修正项,这个误差项使得长期利率与短期利率反复出现分歧,因此,我们还不清楚Kozicki和Tinsley(2001)的批评在使用VECM方法时是否仍然很重要。

宽泛数据集合上的实证结果有“多种可能”的倾向。然而,就整体而言,如果我们不得不对根据大量证据来得出结论,那么我们认为,对5年及以下的期限而言,EH通常表现得非常漂亮,但极端扰动时期要除外(如美国货币基础控制时期1979—1982),这会导致状态变化和严重的比索问题。当然,历史上的名义收益(和收益率差)都非常大。在收益率小且相当平稳的时期中,时变期限溢价对收益率的变化有着很大的贡献。

22.6 小结

* 使用长期利率和完全预见长期利率的方差边界检验面临非平稳性问题,使用完全预见收益率差和实际收益率差可以克服这个问题,因为收益率差是平稳变量。

* 如果长期利率和短期利率是I(1)序列,那么EH意味着(长短期)收益率差是平稳序列,收益率差是未来利率变化的Granger原因。平稳变量的使用意味着我们可以使用标准检验程序,尽管小样本问题仍然存在。

* Var方法在对实际收益率差 S_t 和完全预见收益率差 S_t'（即短期利率未来变化的预测）进行比较的基础上给出了EH的统计检验。

* 在EH为真的零假设下，Var方法所给出的“度量”包括相关系数检验 $\rho(S_t',S_t)=1$、方差比检验 $SDR=\sigma(S_t')/\sigma(S_t)=1$ 和对Var参数的交叉方程限制检验。

* 使用长期利率变化回归和短期利率变化回归（单方程）及Var方法所得到EH有效性的实证结果有多种可能。除美国数据外，这些检验中的绝大多数都支持EH，尽管在扰动期上时变期限溢价似乎很重要（如美国1979—1982）。

第23章

随机折现因子与仿射期限结构模型

目的

* 说明随机折现因子(SDF)方法如何导出期限结构方程,期限结构方程包含依赖于条件方差的风险溢价。

* 说明如何使用仿射模型来解释收益率曲线的形状和运动。

当债券持有期小于债券期限时,债券就成为一种风险资产,因为债券的重售价值不确定。与股票不一样,债券(通常)有固定的到期时间,以固定的面值赎回,所以不需要任何横截性条件。当用随机折现因子(SDF,stochastic Discount Factor)方法来确定长期债券的价格或(即期)收益率时,固定的终端价值就变得非常有用。

23.1 SDF 模型

正如我们即将看到的那样,SDF 方法在长期债券中引入了一种特殊的风险溢价。长期利率不仅与 EH 一样依赖于预期未来短期利率,还依赖于随条件协方差变化的风险溢价,这里的条件协方差是 SDF 与不同期限的长期利率序列之间的协方差。根据原理,这些条件协方差可以运用 GARCH - M(GARCH - in - mean)方法和可观察因子(如消费增长)来进行估计,但这种方法在实证研究中并没有得到广泛应用。

为了使 SDF 模型易于使用,我们经常假设仿射(affine)或线性结构存在,债券价格线性地依赖于潜在的单因子 z_t 或多因子 z_{it}。在单因子模型中,尽管 z_t 在初始时刻是无法观察的因子,在某些假设下还是可以证明,z_t 线性地依赖于可观察的短期利率。这使得我们能导出收益率曲线形状和位置的显性形式,因此也能给出利率期限结构的 SDF - 仿射模型(SDF - affine model)。以下讨论离散形式的仿射模型,但也可以在连续时间情形中应用格子法(lattice approach)来导出这些仿射模型(简单介绍参见 Cuthbertson 和 Nitzshe(2001b))。EH 和 SDF - 仿射方法之间的关键差异在于,后者在收益率曲线上嵌入了与无套利机会一致的风险溢价。

为了发展利率的 SDF 模型,我们以通常的均衡条件(FOC)作为开始:

$$E_t[M_{t+1}(1+H_{n,t+1})]=1 \tag{1}$$

其中,

$$(1+H_{n,t+1})\equiv\frac{P_{n-1,t+1}}{P_{n,t}} \tag{2}$$

$$P_{n,t}=e^{-nR_{n,t}}$$

$$R_{n,t} = -\frac{1}{n}p_{n,t} \tag{3}$$

$(1+H_{n,t+1})$是到期时支付1美元的 n 期零息债券的持有期收益率，$R_{n,t}$是债券的即期收益率。取对数并运用 $h_{n,t+1}=\ln(P_{n-1,t+1}/P_{n,t})$ 就得到

$$h_{n,t+1} = p_{n-1,t+1} - p_{n,t} = nR_{n,t} - (n-1)R_{n-1,t+1} \tag{4}$$

其中，$p_{n,t}=\ln P_{n,t}$。当 $n=1$ 时，债券为无风险债券，根据(1)，有

$$(1+r_t)E_t M_{t+1} = 1 \tag{5}$$

我们在附录 I 中将证明，在 $P_{n,t}$和 M_{t+1}服从联合正态分布的条件下 FOC(1)给出了大家熟知的均衡无套利条件：

$$E_t(h_{n,t+1}-r_t) + \frac{1}{2}V_t(h_{n,t+1}) = -\mathrm{Cov}_t(m_{t+1},h_{n,t+1}) \tag{5}$$

其中，$m_{t+1}=\ln M_{t+1}$，$V(.)$为条件波动率。左边第二项为 Jensen 效应，“Cov_t”为期限(风险)溢价。在附录 I 中，我们证明了 n 期债券的均衡收益率为

$$R_{n,t} = \frac{1}{n}\sum_{i=0}^{n-1} E_t r_{t+i} - \frac{1}{2n}\sum_{i=0}^{n-1}(n-i-1)^2 V_t(R_{n-i-1,t+i+1}) - \frac{1}{n}\sum_{i=0}^{n-1}(n-i-1)\mathrm{Cov}_t(m_{t+i+1},R_{n-i-1,t+i+1}) \tag{6}$$

右边第一项是(风险中性)“预期假说”，但在 SDF 模型中，长期利率也依赖于条件方差项(即 Jensen 效应)和协方差项，协方差刻画了风险溢价。即使在风险中性 $\mathrm{Cov}_t(\,,\,)=0$ 条件下，SDF 模型也支持无套利条件，但“EH”却不成立，因为 Jesen 效应存在。然而，实证研究却发现后者可以忽略不计。因此，EH(和上一章所讨论的检验方法)假设相对于短期利率的预期运动而言，协方差项可以忽略不计。

名义利率 r_t 由实际利率 rr_t 和预期通货膨胀率 $E_t\pi_{t+1}$来确定，因此(6)给出的收益率曲线形状受后面两个变量和期限溢价(及 Jensen 效应)影响。然而，需要注意的是，(5)和(6)从目前来看没有任何用途，因为到目前为止我们还没有对 SDF 的决定因素进行定义。

对 SDF 模型的检验：可观察因子

检验(5)中 $h_{n,t+1}$所对应的 SDF 模型(在对数正态分布假设下)的直接方法是，假设 m_{t+1}与一组可观察因子 z_{it}之间存在线性关系：

$$m_{t+1} = a + \sum_{i=1}^{m} b_i z_{i,t+1} \tag{7}$$

持有期收益率(HPR)对数值的无套利条件变成

$$E_t(h_{n,t+1}-r_t) + \frac{1}{2}V_t(h_{n,t+1}) = -\sum_{i=1}^{m} b_i\mathrm{Cov}_t(z_{i,t+1},h_{n,t+1}) \tag{8}$$

可以任意选择可观察因子，但可观察因子可能包括消费增长(CCAPM 模型)、通货膨胀率和长短期利率差的变化。可能影响债券风险的其他变量包括汇率变化、货币供给和实际产出增长率等“通常的猜测”。困难在于如何对条件协方差进行建模，人们试图使用多变量 GARCH 模型来解决这一困难。例如，如果使用两个长期利率($n=1,2$)，那么

$$x_{t+1} = (h_{1,t+1}-r_t,\ h_{2,t+1}-r_t,\ z_{1t},\ 2_{2t},\cdots)'$$

$$x_{t+1} = a + Bx_t + D\mathrm{vech}(H_{t+1}) + \varepsilon_{t+1}$$

$$\varepsilon_{t+1}\mid\Omega_t \sim N(0,H_{t+1})$$

考虑 $n=1,2$ 的情形，HPR 方程(8)意味着如下限制条件成立：$e1'B=e2'B=0$，其中$(e1,e2)$是(第一，第二)个分量为1其余分量为零的向量。这可以保证 $h_{n,t+1}-r_t$ 不依赖于自身的滞后值。另外，D 的第一行为$\{0,-1/2,-d_{13}\}$，第二行为$\{-1/2,0,-d_{23}\}$。这不仅保证了方差项进入 $h_{j,t+1}-r_t(j=1,2)$的方程，还保证了 d_{13}和 d_{23}是从条件协方差中选择的。一般而言，困难在于如何限制 D 的“规模”，因为根据原理，D 包含 x_{t+1}中所有变量的方差和协方差。尽管可以对 D 施加特殊

限制（如 ε_i 与 ε_j 之间的相关系数为常数），较大的待估参数数目（如果期限个数超过3，这种情形尤其突出）意味着模型在通常情形中无法使用。这是多因子 SDF 模型的一个问题（如何同时估计所有的条件方差）。如果人们试图使用随机波动率模型（模型在波动率过程中增加误差项），那么这个问题会变得更加复杂。

注意，如果使用关于消费的幂效用函数和名义利率，那么我们有

$$m_{t+1} = \psi - (\text{crra})\Delta c_{t+1} - \pi_{t+1} \tag{9}$$

其中，π_{t+1} = 通货膨胀率。分析中一般只出现这两个变量。现有文献很少应用可观察因子来进行直接检验，可能是因为估计很困难，也可能是因为消费和 HPR 的条件方差非常小。然而，一个重要例外是 Balfoussia 和 Wickens（2003），他们使用美国 1970—1998 的月度数据来进行实证研究。他们使用的是上述多变量均值 GARCH（MGM，GARCH - M）模型，GARHCH 方程成“对角形”，目的是为了使每个条件方差只依赖于自身的历史值和历史异常值。就 CCAPM 模型而言，我们有 $z_{1t} = \Delta c_{t+1}$ 和 $z_{2t} = \pi_{t+1}$（参见方程（9）），n 期债券的 HPR 为

$$E_t(h_{n,t+1} - R_t) + \frac{1}{2}V_t(h_{n,t+1}) = (crra)\text{Cov}_t(h_{n,t+1}, \Delta c_{t+1}) + \text{Cov}_t(h_{n,t+1}, \pi_{t+1})$$

因此，$\text{Cov}_t(h_{n,t+1}, \Delta c_{t+1})$ 的参数为相对风险规避系数，$\text{Cov}_t(h_{n,t+1}, \pi_{t+1})$ 的系数应为 1。对不同期限而言（即对不同的 n 而言，在 $h_{n,t+1} - r_t$ 所对应的不同方程中），这些常系数为模型检验提供了另外一种方法。注意，尽管 ccra 在不同方程中是常数，期限溢价随着期限（和时间）的不同而不同，因为对不同期限而言，估计的条件协方差项（来自于 GARCH 过程）不同。不幸的是，美国数据表明 CCAPM 模型表现效果不是很好，因为所估计的 ccra 为 36（在任何一个 HPR 方程中都不显著）。$\text{Cov}_t(h_{n,t+1}, \pi_{t+1})$ 项的系数也不等于 1，在不同的方程中也不是常数。当增加宏观经济因子来导出更为一般的 SDF 模型时（见方程（7a）和（8）），条件协方差项的系数就不会受到限制。现在，$\text{Cov}_t(h_{n,t+1}, \pi_{t+1})$ 项在条件平均收益 $E_t h_{n,t+1} - r_t$ 方程中是统计显著的，但这些协方差项的系数相对于不同的期限而言并不是一个常数。后一结果也适用于消费的协方差项（模型使用 5 个不同期限债券的 HPR）。

使用方程（8），由 MGM 模型来估计时变协方差项，于是我们可以对如下方程进行估计：

$$h_{n,t+1} = \alpha + \beta r_t + \gamma\phi_{n,t} + \varepsilon_{n,t+1}$$

或等价地，对长期利率变化方程进行估计

$$(n-1)(R_{t+1}^{(n-1)} - R_t^{(n)}) = \delta + \mu(R_t^{(n)} - r_t) + \upsilon\phi_{n,t} - \varepsilon_{n,t+1}$$

其中，$\phi_{n,t}$ 是（8）中协方差项的和。根据 EH，$\alpha = 0, \beta = 1$ 或 $\delta = 0, \mu = 1$。如果 MGM 模型给出的是时变风险溢价的估计值，那么我们期望 $\gamma = 1$ 或 $\upsilon = -1$。因为 $h_{n,t+1}$ 和 r_t 近似为单位根过程，所以当假定 $\phi_{n,t}$ 是 I(0) 过程（即 $\hat{\beta}$ 是很好的一致估计量）时，在 HPR 方程中省略期限溢价 $\phi_{n,t}$ 后所得到的估计值 $\hat{\beta}$ 不会有太大的偏差。另一方面，长期利率变化方程中的变量是 I(0) 序列，因此省略 $\phi_{n,t}$ 可能会导致 μ 估计值有较大的偏差。不幸的是，在后面一个方程中包括 $\phi_{n,t}$ 的 MGM 估计值将导致 μ 的估计值进一步偏离 +1（对绝大多数期限而言），因此，美国数据所证实的 EH 不成立并不是因为使用了 MGM 模型和 SDF 方法。Balfoussia 和 Wickens（2003）已经证实，超额 HPR 与条件协方差（使用带特殊参数的 MGM 进行建模）之间的关系在统计上并不能充分地给予确定，对不同期限的美国 HPR 而言，时变期限溢价的解释也不稳定。正如在前面章节中所注意到的那样，在将利率高度波动时期（1979—1982）包含进来后，MGM 模型可能出现“异常值”敏感问题。在上述 SDF 方法中，我们使用了可观察因子，其中的一个问题是，我们在实证研究中究竟应该使用哪一个因子，因为我们的分析几乎没有对其进行说明（唯一的例外是，通货膨胀因子是一种相当直观的选择）。一种替代方法是使用无法观察的因子，而这些因子能被嵌入到潜在的仿射因子模型中。

23.2　单因子彷射模型

获得期限溢价测度的一种方法是使用期限结构的彷射模型，Vasicek(1977)和Cox、Ingersoll和Ross(1985)都是SDF模型的特殊情形，在SDF模型中持有期收益率与因子之间的条件协方差是因子本身的线性函数。单因子Vasicek模型意味着收益率曲线的形状和风险溢价只依赖于期限，收益率曲线的形状在整个时期中固定不变。单因子Cox－Ingersoll－Ross模型也意味着收益率曲线的形状固定，但允许风险溢价随着时间的演变而发生变化，这种现象之所以出现是因为短期利率发生了变化。多因子彷射模型能产生收益率曲线形状更大的弹性(参见Piazzesi(2002)的综述)。

简单且易于使用的假设是：债券价格的函数形式是一些(到目前为止还没有定义的)因子z_t的线性函数(我们就使用这种形式)

$$p_{n,t} = -[A_n + B_n z_t] \tag{10}$$

从当前时刻来看，方程(10)完全是任意假设的。它说明，当单因子z_t发生变化且z_t任何变化对$p_{n,t}$所产生的定量影响只依赖于债券期限时，所有期限(n)的债券仅仅沿着时间发生移动。

想要说明的是，彷射(线性)假设加上无套利SDF条件(1)所得出的结果与基于可观察收益率曲线的典型事实是一致。这些典型事实包括：

(1) 通常观察到的收益率曲线是向上倾斜的匀滑曲线，在较长的期限后变为“扁平”形状。

(2) 当收益率曲线向上或向下移动时，它在绝大多数地方保持了向上倾斜的形状，从宽泛意义上讲(但不准确)，不同期限收益率移动的幅度相同(即收益率曲线平行移动)。

(3) 偶尔会观察到收益率曲线向下倾斜。

(4) 偶尔会出现如下情形：向上倾斜的收益率曲线在期限中间具有“驼峰形状”。

如下步骤可以用来对这些彷射模型进行求解：

(a) 首先假设$p_{n,t}$是z_t的线性函数，系数分别为A_n和B_n；

(b) 假设刻画m_{t+1}与z_t之间关系的“合理”线性时间序列模型(time－series model，TSM)能估计出来(通常是用Kalman滤波方法)；

(c) 推导无套利SDF模型所隐含的A_n和B_n的限制条件；

(d) 将A_n和B_n的这些表示代入$p_{n,t}$的线性方程就得到$p_{n,t}$和$R_{n,t} = (-1/n)p_{n,t}$的无套利均衡条件(对任意期限n而言)；

(e) 考察均衡的“理论”收益率曲线是否与上述典型事实一致。

下面给出解释m_{t+1}和z_t联合行为的两个广为人知的时间序列(TSM)模型。

A. Cox－Ingersoll－Ross模型(CIR)

$$-m_{t+1} = z_t + \lambda e_{t+1} \tag{11a}$$

$$(z_{t+1} - \mu) = \theta(z_t - \mu) + e_{t+1} \tag{11b}$$

其中，$e_{t+1} = \sigma\sqrt{z_t}\varepsilon_{t+1}$，$\varepsilon_{t+1} \sim iid(0,1)$，因此$V_t(e_{t+1}) = \sigma^2 z_t$。

B. Vasicek模型

方程(11a)＋(11b)成立，但e_t同方差，即$e_{t+1} = \sigma^2\varepsilon_{t+1}$。

CIR模型的解

使用无套利方程(1)所对应的上述时间序列模型，通过冗长的计算(参见附录II)就可以得到A_n和B_n的解，因此，我们也能得到$p_{n,t}$、和$R_{n,t}$期限溢价的解。A_n和B_n是参数$(\mu, \phi, \lambda, \sigma^2, n)$的函数。可以看出，短期利率$r_t$是$z_t$的精确线性函数，这使得我们在任意时点都可以用能观察

的 r_t 来代替不能观察的 z_t。(根据原理,用 r_t 来代替 z_t 通常是不可能的。)在附录 II 中,我们证明 Cox - Ingersoll - Ross 模型(CIR)所对应的形式为

$$r_t = \left(1 - \frac{1}{2}\lambda^2\sigma^2\right)z_t \tag{12}$$

$$E_t(h_{n,t+1} - r_t) = \left(\frac{1}{2}B_{n-1}^2\sigma^2 + \lambda B_{n-1}\sigma^2\right)z_t \tag{13}$$

方程(13)的第一项是 Jensen 效应,第二项是风险溢价

$$rp_t = -\mathrm{Cov}_t(m_{t+1}, p_{n-1,t+1}) = \lambda B_{n-1}\sigma^2 z_t \tag{14}$$

最后,可以证明不同期限的长期利率线性地依赖于单因子 z_t,因此,长期利率线性地依赖于短期利率 r_t:

$$\begin{aligned} R_{n,t} &= -\frac{1}{n}p_{n,t} = \frac{1}{n}(A_n + B_n z_t) \\ &= \frac{A_n}{n} + \frac{B_n}{n\left(1 - \frac{1}{2}\lambda^2\sigma^2\right)}r_t \end{aligned} \tag{15}$$

Vasicek 模型的解

我们只给出收益率曲线和期限(风险)溢价:

$$R_{n,t} = \frac{1}{n}\left(A_n - \frac{1}{2}\lambda^2\sigma^2 B_n\right) + \left(\frac{B_n}{n}\right)r_t \tag{16}$$

$$rp = \lambda\frac{(1 - \theta^{n-1})\sigma^2}{(1 - \theta)} \tag{17}$$

A_n 和 B_n 再次依赖于参数($\lambda, \mu, \theta, n, \sigma^2$)。从方程(15)和(16)很容易看出,在 CIR 模型和 Vascicek 模型中,收益率曲线的形状依赖于 n,当短期利率 r_t 变化时,曲线向上或向下移动。CIR 模型和 Vascicek 模型都意味着收益率曲线的形状关于时间是常数,但我们在实践中知道收益率曲线的倾斜方向会发生变化(有时向上倾斜,有时向下倾斜)。因此,这些“单因子”仿射模型不能拟合典型事实。另外,在 Vascicek 模型中,风险溢价不随时间的演变而发生改变,但风险溢价只取决于期限(这是一个带限制条件的结果)。

我们有两种方法来给出能拟合更多“典型事实”且易于使用的 SDF 模型。最为明显的扩展方法是在仿射模型中考虑多个因子。即使只选取两个因子,根据因子的不同运动,收益率曲线的形状也会发生变化。第二种扩展方法是选用多因子模型(为了易于使用,一般限制为三个因子),并增加诸如通货膨胀和实际活动等可观察的宏观经济因子。

23.3 多因子仿射模型

现在,我们有多个因子 z_t,当 z 为向量时,上述方程同样成立(参数用矩阵形式来进表示)。这些模型的复杂性和待估参数的数目会迅速增加,因此我们经常要对模型施加一些简化条件。例如,如果方程(11b)中的调整矩阵 θ 是对角矩阵、向量误差 $e_{t+1} = \sum\sqrt{S_t}\varepsilon_{t+1}$中 $\sum$ 独立于 θ、S_t($= \mathrm{diag}z_{it}$)是对角矩阵,那么方程(10)对 z_{it}而言具有可加性:

$$E_t h_{n,t+1} - r_t = \frac{1}{2}\sum_{i=1}^{m}B_{i,n-1}^2\sigma_i^2 z_{it} + \sum_{i=1}^{m}\lambda_i B_{i,n-1}\sigma_i^2 z_{it} \tag{18}$$

最后一项是风险溢价。然而,短期利率现在依赖于 z_{it}的一个线性组合,我们不再用可观察的 r_t 来“替换”z_{it}。如果收益率的个数大于 n(包括短期利率),那么 n 个因子不是 m 个收益率的唯一线性组合,我们需要使用主成分形式在 n 个无法观察的因子中“排列”m 个可观察的收益率($m < n$)。然而,因为存在更多的因子,所以收益率曲线的形状更加富有弹性。一般而言,两个或三个

因子就能充分逼近所能观察到的收益率曲线的不同形状。如果收益率的个数比因子个数还要多，那么这个模型是无法观测的因子模型。

实证结果

Gong 和 Remolona(1997)及 Dai 和 Singleton(2003)对双因子和三因子 CIR 模型进行了估计，而 Jagadeesh 和 Pennacchi(1996)对单因子和双因子 Vasicek 模型进行了估计(使用欧洲美元期货合约(流动性很强的市场))。在多于一个因子的模型中，“短期利率”不能“替换”因子，从而估计变得非常复杂。双因子和三因子模型一般会拒绝假设检验中的单因子模型，但他们经常保留了一些不合理的效应(例如，在一半的时间中状态变量出现异常值，短期利率有时为负数)。Ang 和 Piazzesi(2003)增加了“通货膨胀”和“实际产出”两个宏观经济因子，而这些因子是从一个更大的宏观经济变量集合中使用主成分分析方法来获得的。无法观察的三因子模型包括这两个宏观经济因子，三因子模型已用于 1、6、12、36、60 月债券收益率的研究。在应用过程中，还假设所有因子之间是相互独立的。这些可观察的宏观经济因子和 $z_{it}(i=1,2,3)$ 会对短期利率和收益率曲线的形状产生影响。这些仿射方法是否充分地拟合了所有的“典型事实”，这是一个仍需考察的问题。

仿射方法也适用于指数债券，指数债券的名义长期利率 $R_{n,t}^{(Now)}$ 和实际长期利率 $R_{n,t}^{(real)}$ 都具有如下形式的线性期限结构方程：

$$R_{n,t}^{(j)}=\frac{1}{n}\left[\sum_{i=0}^{n-1}\left(E_t r_{t+i}^{(j)}+E_t \phi_{t+i}^{(j)}\right)\right] \tag{19}$$

其中，$j=Nom$ 或 $real$，ϕ 表示期限溢价和 Jensen 项。使用 Fisher 方程 $r_t^{Nom}=r_t^{real}+E_t\pi_{t+1}$($\pi_{t+1}$ 为通货膨胀率)就得到

$$\frac{1}{n}\sum_{i=0}^{n-1}E_t\pi_{t+i}=R_{n,t}^{(Nom)}-R_{n,t}^{(real)}-\frac{1}{n}\sum_{i=0}^{n-1}E_t\phi_{t+i}^{\pi} \tag{20}$$

预期(平均)通货膨胀率等于名义收益率和实际收益率之差减去通货膨胀风险溢价 $\phi^{\pi}\triangleq\phi^{Nom}-\phi^{real}$。例如，Remolona、Wickens 和 Gong(1998)应用英国上世纪 90 年代的数据发现，通货膨胀风险溢价具有时变特征，且平均水平在 1% 左右。

23.4　小结

* SDF 模型(对数正态分布假设)意味着超额 HPR 不是常数，但依赖于(可能的)时变条件协方差(和 Jensen 效应)。

* SDF 模型是比 EH(风险中性假设)更为一般的模型。在 SDF 模型中，长期利率不仅依赖于预期未来短期利率，还依赖于 SDF 与 n 期债券收益率之间的协方差。EH 的实证检验假设这些协方差(和 Jensen 效应)，相对于预期未来短期利率的变化而言，可以忽略不计。

* 期限结构的仿射模型是 SDF 模型的一种特殊形式，它假设随机折现因子(m_{t+1})的对数是一组因子 z_{it} 的线性函数。仿射模型通常能给出长期债券收益率 $R_{n,t}$ 的显式闭型解，我们由此可以推断收益率曲线的形状和位置。除了一些简单情形外，仿射模型包含无法观察的因子，因而需要复杂的估计技巧。

* 在使用可观察协方差的实证研究中，超额 HPR 的 SDF 模型没有得到广泛考察，可能原因是，“大型”GARCH - M 模型的估计非常困难。实证结果表明，仿射模型(通常选用双因子或三因子)已经能解释收益率曲线的一些特征，但不能解释它的全部特征。

附录 I：期限结构 SDF 模型中的数学

下面叙述的离散版本随机折现因子（SDF）模型基本上来自于 Smith 和 Wickens（2002）。在这一节里，我们建立不同期限价格和收益率的无套利方程及持有期收益率的无套利方程。这些方程中的风险溢价取决于随机折现因子与价格或收益率之间的协方差。对 SDF 模型而言，无套利条件为

$$E_t[M_{t+1}(1-h_{n,t+1})]=1 \tag{A1}$$

或

$$P_{n,t}=E_t[M_{t+1}P_{n-1,t+1}] \tag{A2}$$

当 $n=1$ 时，债券是无风险债券，因此

$$(1+r_t)E_t[M_{t+1}]=1 \tag{A3}$$

根据无风险利率的定义，上述方程定义了 $E_t(M_{t+1})=1/(1+r_t)$。如果 $P_{n,t}$ 和 M_{t+1} 服从联合对数正态分布，那么根据（A2），有

$$\begin{aligned}p_{n,t}&=E_t(m_{t+1}+p_{n-1,t+1})+\frac{1}{2}V_t(m_{t+1}+p_{n-1,t+1})\\&=E_tm_{t+1}+E_tp_{n-1,t+1}\\&\quad+\frac{1}{2}[V_t(m_{t+1})+V_t(p_{n-1,t+1})+2\mathrm{Cov}_t(m_{t+1},p_{n-1,t+1})]\end{aligned} \tag{A4}$$

将 $p_{0,t}=0$ 代入（A4）就得到单期债券的价格：

$$p_{1t}=E_t(m_{t+1})+\frac{1}{2}V_t(m_{t+1}) \tag{A5}$$

（A4）减去（A3）就得到价格（也就是超额 HPR）的无套利方程

$$E_t(p_{n-1,t+1})-p_{n,t}+\frac{1}{2}V_t(p_{n-1,t+1})=-\mathrm{Cov}_t(m_{t+1},p_{n-1,t+1}) \tag{A6a}$$

$$E_t(h_{n,t+1}-r_t)+\frac{1}{2}V_t(p_{n-1,t+1})=-\mathrm{Cov}_t(m_{t+1},p_{n-1,t+1}) \tag{A6b}$$

使用 $\ln P_{n,t}=-n\ln(1+R_{n,t})\approx-nR_{n,t}$，（A6）可以用收益率的形式来进行表示。

$$\begin{aligned}&-(n-1)E_t(R_{n-1,t+1})+nR_{n,t}-r_t+\frac{(n-1)^2}{n}V_t(R_{n-1,t+1})\\&=(n-1)\mathrm{Cov}_t(m_{t+1},R_{n-1,t+1})\end{aligned} \tag{A7}$$

方程（A7）就是前面为 EH 命名的“长期利率变化”方程。EH 假设风险溢价为常数，而 SDF 模型中的风险溢价是 $\mathrm{Cov}_t(\,,\,)$ 项，$V_t(\cdot)$ 项是 Jensen 不等式项（JIT）。

将 $h_{n,t}=-(n-1)R_{n-1,t+1}+nR_{n,t}$ 进一步代入（A7）就得到预期超额持有期收益率的无套利方程：

$$E_t(h_{n,t+1}-r_t)+\frac{1}{2}V_t(h_{n,t+1})=-\mathrm{Cov}_t(m_{t+1},h_{n,t+1}) \tag{A8}$$

方程（A6）～（A8）是联合对数正态分布假设下无套利方程（A1）的替代表示。协方差项是风险溢价，出现方差项的原因是 Jensen 效应存在。最后，方程（A7）向前迭代就得到用预期未来短期利率表示的长期利率（加上期限溢价（和 Jensen 效应））。

$$\begin{aligned}R_{n,t}=&\frac{1}{n}\sum_{i=0}^{n-1}E_tr_{t+i}-\frac{1}{2n}\sum_{i=0}^{n-1}(n-i-1)^2V_t(R_{n-i-1,t+i+1})\\&-\frac{1}{n}\sum_{i=0}^{n-1}(n-i-1)\mathrm{Cov}_t(m_{t+i+1},R_{n-i-1,t+i+1})\end{aligned} \tag{A9}$$

方程（A9）指出了一个显而易见的事实，纯预期假说是 SDF 模型的一种特殊情形（在对数正态分布的假设下），在纯预期假说中风险溢价为零，Jensen 效应不存在。当然，在对随机折现因子

m_{t+i+1}和方差项进行建模之前，我们仍然无法对模型进行检验。

附录 II：单因子仿射模型

导出收益率曲线形状显式方程的一种方法是：假设 m_{t+1} 与单因子 z_t 线性相关（在下面的部分中可以看到，z_t 有时可以用可观察的短期利率来进行度量）。这些仿射模型假设 $p_{n,t}$是因子 z_t 的线性函数。

$$p_{n,t}=-[A_n+B_nz_t] \tag{A10}$$

因此，

$$R_{n,t}=-\frac{1}{n}p_{n,t}=\frac{A_n}{n}+\frac{B_n}{n}z_t \tag{A11}$$

且

$$r_t=-p_{1,t}=A_1+B_1z_t \tag{A12}$$

现在，我们来导出 A_n 和 B_n 满足 SDF 模型无套利条件（$A1$）的（递归）限制条件，紧接着将（$A6b$）超额 HPR 方程中的 $E_th_{n,t+1}-r_t$ 进行比较来得到用 A_n 和 B_n 表示的风险溢价。m_{t+1}和 z_t 的一个合理且易于使用的时间序列模型（TSM）是著名的 Cox、Ingersoll 和 Ross（1985）（CIR）模型。

$$-m_{t+1}=z_t+\lambda e_{t+1} \tag{A13a}$$

$$(z_{t+1}-\mu)=\theta(z_t-\mu)+e_{t+1} \tag{A13b}$$

其中，$e_{t+1}=\sigma\sqrt{z_t}\varepsilon_{t+1}$，$\varepsilon_{t+1}\sim iid(0,1)$，$V_t(e_{t+1})=\sigma^2z_t$。在这个模型中，条件均值 m_{t+1}等于 z_t（参见（A13a））。假设因子 z_t 是围绕长期均值 μ 运动的自回归序列，均值反转速度由 θ 来决定。现在来考虑方程（A4）右边的项，导出 $p_{n,t}$的公式。由（A13a），（A13b）和（A10）就得到，

$$\begin{aligned}E_t(m_{t+1}+p_{n-1,t+1})&=-[z_t+A_{n-1}+B_{n-1}E_tz_{t+1}]\\&=-[z_t+A_{n-1}+B_{n-1}(\mu(1-\theta)+\theta z_t)]\end{aligned} \tag{A14}$$

且

$$V_t(m_{t+1}+p_{n-1,t+1})=V_t(\lambda e_{t+1}+B_{n-1}e_{t+1})=(\lambda+B_{n-1})^2\sigma^2z_t \tag{A15}$$

将（A14）和（A15）代入无套利方程（A4）并使用 $p_{n,t}$的"仿射方程"（A10）就得到一个更为复杂的方程：

$$\begin{aligned}p_{n,t}&=-[A_n+B_nz_t]\\&=-[(1+\theta B_{n-1})z_t+A_{n-1}+B_{n-1}\mu(1-\theta)]+\frac{1}{2}(\lambda+B_{n-1})^2\sigma^2z_t\end{aligned} \tag{A16}$$

令方程左右两边的系数相等，就得到 A_n 和 B_n 的递归公式：

$$A_n=A_{n-1}+B_{n-1}\mu(1-\theta) \tag{A17a}$$

$$B_n=1+\theta B_{n-1}-\frac{1}{2}(\lambda+B_{n-1})^2\sigma^2 \tag{A17b}$$

因为 $\ln P_{0,t}=p_{0,t}=0$，所以 $A_0=B_0=0$ 为递归方程（A17）的初值条件。当 $n=1$ 时，我们有

$$B_1=1-\frac{1}{2}\lambda^2\sigma^2\quad 且\quad A_1=0 \tag{A18}$$

目的是为了使下式成立

$$r_t=-p_{1,t}=A_t+B_1z_t=(1-\frac{1}{2}\lambda^2\sigma^2)z_t \tag{A19}$$

因此，短期利率与 z_t 线性相关，我们可以将 $p_{n,t}$和 $R_{n,t}$的方程（A10）和（A11）中无法观察的 z_t 用可以观察的短期利率 r_t 来"替代"。根据（对数）价格可以定义持有期收益率，由（A16）和（A19）可以导出

$$E_t(h_{n,t+1}-r_t)\equiv E_t(p_{n-1,t+1})-p_{n,t}-p_{1,t}$$

$$= \frac{1}{2} P_{n-1}^{2} \sigma^{2} z_{t} + \lambda B_{n-1} \sigma^{2} z_{t} \tag{A20}$$

由（A6b）有

$$E_{t} h_{n,t+1} - r_{t} = -\frac{1}{2} V_{t}(p_{n-1,t+1}) - \mathrm{Cov}_{t}(m_{t+1}, p_{n-1,t+1}) \tag{A6b}$$

比较（A20）和（A6b）就可以看出：第一项是 Jensen 效应，第二项是风险溢价 $-\mathrm{Cov}_{t}(m_{t+1}, p_{n+1,t+1})$，两项都依赖于 z_t。值得注意的是，如果 m_{t+1} 不是随机变量（即 $\lambda = 0$），那么风险溢价为零。在（A2）中，我们再次可以用 r_t 来“代替”z_t。由（A19）可以看出，超额 HPR 和风险溢价一样线性依赖于 r_t。在一个因子驱动的“线性世界”中，这个结论并不奇怪，后者是（对数）价格、长期利率和期限溢价的短暂驱动变量（只有后面一个结果毫无直觉）。Vasicek 模型可以用上述方法来进行精确求解，但我们在这里不再赘述。

第 24 章

外汇市场

目的

* 对各种汇率制度进行简要概述。

* 说明价格竞争力如何用购买力平价(PPP)和一价定律(LOOP)来进行表述。

* 说明无风险套利如何决定远期汇率。这就是抵补利率平价(CIP)。

* 说明如果(风险中性投机者)消除了风险套利机会,那么非抵补利率平价(UIP)就会产生。

* 分析 CIP、UIP 和远期利率无偏性(FRU)之间的关系及 UIP、PPP 和实际利率平价(RIP)之间的关系。

24.1 汇率制度

汇率的行为,尤其是对那些有着大量国际贸易的小型开放经济而言,多年以来一直是宏观经济政策争论的中心。毫无疑问,就采用什么样的最佳汇率体系而言,经济学家的观点每年都在改变。因此,简略地勾画主要问题是值得的。

第二次世界大战后,绝大多数货币都采用布雷顿森林体系,实行“固定但可以调整的”汇率制度。在资本流量小且经常遭到政府管制的时期中,这些国家强调的重点是价格竞争力。那些通货膨胀率比贸易伙伴高的国家最初被允许从国际货币基金组织(IMF)借款来为财政赤字融资。如果贸易账户出现了根本性不平衡,那么通过协商,允许贸易赤字国将汇率固定在一个新的更低水平上。在货币贬值后,IMF 通常也会采取一些严厉的措施(例如削减公共开支)来保证实体资源(劳动力和资本)的流入有助于出口增长和进口替代。这一体系多年来运行相对良好,成功地避免了 20 世纪 30 年代贸易保护主义的高关税和配额的再现。

美元是布雷顿森林体系的锚货币,它最初与黄金挂钩的固定价格是每盎司 35 美元。在 20 世纪 60 年代中期,这一体系的问题开始凸现出来。贸易赤字国不能说服贸易盈余国通过重新估值盈余国货币的方式来减缓竞争力问题。贸易国之间的非对称调整过程意味着赤字国的货币必须贬值。大幅贬值的可能不仅使投机者能进行“单向赌博”,而且鼓励投机者对那些经常账户有明显不平衡的国家进行投机性攻击,即使理性观点认为这种不平衡只是暂时的,投机者也会这样做。美国持有大量的经常账户赤字,这就增加了第三世界国家的美元持有量(这正是美国“霸权”的一种体现)。事实上,若按每盎司黄金 35 美元的“官方价格”计算,美元的外部持有量已超过了诺克斯堡(美国联邦政府黄金贮存地)的黄金总量。这就是说,在当时的官方价格下,

黄金与美元的自由兑换已经不可能。两个层次的黄金市场开始发展起来(黄金的市场价格远远高于官方价格),美国官方当局最终中止了美元与黄金的固定比例兑换。在20世纪70年代早期,布雷顿森林体系面临的压力日益增加,因为当时的国际资本更具流动性,国家之间的通货膨胀率差异不断增加,这就导致经常账户出现了大量的赤字和盈余。到了1972年,绝大多数工业国家实际上已经放弃了布雷顿森林体系,实行自由浮动的汇率政策。

向浮动汇率制转变的部分原因是受到了货币经济学家们的影响。他们认为,国内货币供给的良好控制能保证通货膨胀率到达合意水平,能使汇率在合意的轨道上运行。此外,理性投资者的投机活动能确保实际汇率(从而价格竞争力)的大幅持续性波动可以通过可信货币政策的宣布(通常以货币供给目标的相应方式来进行)来加以避免。

到20世纪70年代后期,Dornbusch(1976)的学术论文认为,如果外汇交易商都是理性的且商品价格都是“粘性的”,那么就会发生汇率的过度反应。矛盾的货币政策只会导致价格竞争力在较长时期中趋于减弱,实际贸易、产出和就业出现明显萎缩。尽管从长期均衡来看,经济会实现充分就业和更低的通货膨胀率,但与假定价格具有“弹性”的早期货币模型相比,在Dornbusch模型中,过渡时期产出的减少量要大得多。(我们将在第27章考察由汇率决定的货币模型。)

20世纪70年代名义汇率和实际汇率反复无常的波动致使许多欧洲国家考虑回归管制更多的汇率制度,20世纪80年代早期开始的汇率机制(ERM)就是这种思潮的体现。加入ERM的欧洲国家一致同意将双边汇率维持在以平价汇率为中心的某个区间内,区间范围可宽(±6%)可窄(±2.25%)。德国马克(DM)成为非官方的锚货币。ERM用德国货币政策来取代国内货币政策,以此为工具来对付国内通货膨胀。因为ERM国家的货币与德国马克(DM)之间存在固定的兑换比率,所以高通货膨胀清晰地发出了一个信号,这些国家必须降低通货膨胀率来与德国保持一致,否则失业就会在高通货膨胀国家中发生。由此看来,这是降低通货膨胀的一个“痛苦机制”。ERM允许成员国在基础性失衡的情形下重新调整自己的(中心)汇率。然而,如果是因为随机的投机性攻击致使一国货币到达汇率范围的底部,那么体系内所有中央银行都会在外汇市场通过协调干预来支持这种弱势货币。

ERM在20世纪80年代在降低通货膨胀和汇率波动率上所取得的成功使得G10开始考虑采用协调干预政策(即广场和罗浮宫协议)来减缓对本国货币“不利的”持续性变动。协调干预政策在1983—1985年期间美元的“持续”上升中得到了体现,这种上升似乎与经济基础的变化完全无关。最近,一些经济学家建议根据ERM规则来设定更为正式的货币区域和主要货币的浮动范围。

然而,在20世纪90年代早期,ERM遇到了极大的压力。日益增长的资本流动性和ERM国家对外汇管制的废除,大大地便利了1992年9月16日(著名的黑色星期三)对意大利里拉、英镑和法郎的投机性攻击。英镑和里拉放弃了ERM,允许汇率自由浮动。大约一年以后,面对进一步的货币动乱,绝大多数货币的ERM范围扩大到±15%。欧洲出现货币联盟的原因纷繁复杂,但有一点是毫无疑问的,这些国家想“根除”汇率浮动或对汇率进行半管制所带来的问题。货币联盟正式启动于1991年11月20日在荷兰马斯特里赫特举行的欧盟首脑会议,会议制定出了加入统一货币的“标准”和时间表。

1999年1月,11个欧盟国家组成了货币联盟,义无反顾地锁定了成员国货币对欧元的汇率,从2002年1月开始,零售交易全部使用欧元的“纸币和硬币”。1999年加入欧元区的国家有奥地利、比利时、芬兰、法国、德国、爱尔兰、意大利、卢森堡、荷兰、葡萄牙和西班牙,希腊在2002年加入了“欧元区”。大体而言,这些国家满足“马斯特里赫特”的通胀和利率目标,但也到达了“创造会计”中所规定的3%的最高预算赤字标准。当然,也有一些国家没有达到负债比率标准(如比利时和意大利的负债-GDP比率比标准大了100%)。因此,马斯特里赫特标准最终只是进入欧元区的“指导方针”。欧盟国家在开始时没有加入欧元区的有英国、丹麦、希腊(已于2002年加入)和瑞典。

欧元作为流通货币在最先加入的11个欧洲国家中使用。它不仅用于商业贸易的计价,也有

许多金融资产（如贷款、股票、债券、远期合约、互换协议和衍生工具）用它来进行标价。欧元成为美元的竞争对手，最终成为重要的交易媒介和储备货币。然而，在与美元的兑换中欧元经历过连续的下跌，从 1999 年 1 月的每欧元 1.17 美元跌至 2000 年 10 月的每欧元 0.83 美元左右。不过，自那以后就开始出现转机，欧元在 2003 年再度回到最初的 1.1 美元的位置。主要“经济体”的名义和实际汇率当然也会进行大幅波动。

在欧洲以外的远东，1997—1999 年也见证了一场巨大的货币动乱，中国台湾、印度尼西亚、马来西亚和日本的银行危机导致这些国家（和地区）对美元的汇率贬值了大约 30% ~40%。外资撤离的原因是国内银行所借的外币“太多”。这些外币随即转化成本地贷款（多数用在房地产投机上），最终变成“不良”贷款。因此，正确的看法是，没有进行套期保值的银行无法归还外币货款的利息和本金，由此触发了国内资本的大量外流。

1998 年，俄罗斯卢布对美元的汇率经历了急速的贬值，也是因为俄罗斯的银行外币借款有违约的倾向。俄罗斯的经济没有实质性的增长，税收收入也在大幅下降（征税系统也处于崩溃的边缘），许多公共部门的工资拖欠超过了六个月。1999 年，巴西“里尔”成为下一个牺牲品。面对投机压力，巴西货币经历了大幅贬值，这使得 IMF 和来自美国的担保贷款都无法根除货币贬值的影响。这一系列事件致使 G10 呼吁“新的经济秩序”或“建构新的金融秩序”，这些都需要 IMF 在试图避免货币危机或减轻对经济的不利影响过程中发挥更为积极的作用。

本书这一部分的主要任务是考察为什么在固定汇率和浮动汇率的选择上会存在这样的困惑，为什么存在这样多的争论，为什么会出现共同的货币区域。有时会出现自相矛盾的情形是，经济学家在设定价格时通常都支持“无约束市场”的观点，但在汇率的设定过程中（也许是经济中的重要“价格”）却存在众多的分歧意见。

24.2　购买力平价与一价定律

一价定律（Law of One Price，LOOP）是同质可交易商品的市场均衡条件，是许多基于经济基本面的汇率模型的构筑基石。它是一种“商品套利”关系。例如，如果单独地应用于国内经济，那么 LOOP 意味着“林肯大陆”（轿车名）在纽约的出售价格和在华盛顿的出售价格应该相等（忽略这两座城市之间的运输成本）。如果在纽约的价格低一些，那么纽约的需求就会较高一些而华盛顿的需求则会低一些。这会导致纽约的价格出现上涨，而华盛顿的价格出现下降，最终使二者相等。事实上，需求转移威胁足以使信息充分的交易者确信两座城市之间的价格相等。LOOP 在不同国家之间也是适用的，唯一不同的是，为了进行比较，必须把其中的一种价格转换成“共同货币”。购买力平价（Purchasing Power Parity，PPP）理论认为 LOOP 对一些价格指数（如批发零售物价指数或消费物价指数 CPI）是适用的，但是，因为我们不讨论指数问题，所以在随后的内容中，你可以认为 LOOP 和 PPP 等价。

绝对购买力平价

如果国内可交易商品是国外商品的完美替代，且商品市场“有效”（即交易成本低、完全信息、价格完全富有弹性、交易无人为限制和政府管制等），那么“中间商”或套利交易者的行动将保证用同种货币表示的价格相等。如果外币价格是 P^* 美元，S 代表汇率（国内货币/国外货币，比如说，英镑/美元），那么国外进口商品的价格用国内货币（如英镑）表示就是 SP^*。国外商品的国内替代生产和市场上的套利交易者将保证国内（英镑）价格 P 等于用国内货币表示的商品进口价格 SP^*：

$$P = SP^* \quad \text{或}\ p = s + p^* \tag{1}$$

$p = \ln p$，其余类似。这里有一个简单的例子。伦敦哈罗斯（Harrods）百货商店在出售“一篮子食品”（也就说，由“固定比例”的商品所组成的篮子，例如，包括馅饼、香槟等），纽约的萨克斯

(Saks)百货也在出售这样的商品。如果萨克斯“一篮子食品”要花200美元，即期汇率是0.6667英镑/美元(或1.5美元/英镑)，那么对美国居民来说，萨克斯“一篮子食品”的价格是133.34英镑。PPP意味着套利交易对哈罗斯“一篮子食品”定价也为133.34英镑，否则在两个国家之间进行运输“一篮子食品”(如在环球信息网上进行购买)就能产生一个近似的无风险获利机会。如果LOOP总是成立，且萨克斯“一篮子食品”的价格上涨到210美元，那么哈罗斯“一篮子食品”的价格也会上涨到140英镑(S仍为0.6667英镑/美元)。

相对购买力平价

相对PPP认为P和SP^*可能不相等，但P的变动与SP^*的变动成比例，因此$P = k(SP^*)$，从而，

$$\Delta p = \Delta s + \Delta p^* \tag{2}$$

PPP也可以视为国际收支平衡表中经常账户的一个平衡条件(国内需求和世界贸易水平给定)。道理很简单，因为如果PPP在任何时期都成立，那么这也就意味着国内居民没有动力将对国外商品的需求转换成对国内替代品的需求(反之亦然)。

实际汇率是价格竞争力的一个测度指标，它是相对对于国外商品(用同一种货币进行计价)而言的国内商品价格。

$$S^r = \frac{P^* S}{P} \tag{3}$$

我们仍然使用前面的例子，最初的实际汇率$S_0^r = (P^* S)/P = 200 \times 0.6667/133.34 = 1$意味着，对英国居民来说，哈罗斯“一篮子食品”和萨克斯“一篮子食品”花费的英镑数相同(当然，对美国而言所花费的美元数也相同)。根据实际汇率的定义，如果PPP成立，那么实际汇率或价格竞争力仍然为常数。但是，实际汇率在现实中会绝对不会是常数，实际汇率在长期内存在“均值反转”效应，这意味着我们很难在一段时期内保持一个稳定的价格竞争力。

如果商品套利是影响汇率的唯一因素，那么汇率满足PPP

$$s = p - p^* \tag{4a}$$

或

$$\Delta s = \Delta p - \Delta p^* \tag{4b}$$

因此，汇率变动是通货膨胀率差异的反映，后者在高通货膨胀国家中经常出现(如20世纪20年代的德国、20世纪80年代的拉丁美洲国家、过渡时期的西欧经济体、1990年后的俄罗斯)。然而，我们期望的结果是，在那些有适度通货膨胀且交易商品众多的发达工业经济中，商品套利可能运行得不是那么完美。因此，在这些经济体中，PPP只在长期内才会成立。

购买力平价的证据

绝对PPP意味着同一商品在不同国家的售卖价格应相等。例如，用美元标价的黄金，不论是在英国、欧元区还是在美国，售卖价格都应相等。当然，绝对PPP对一些可交易商品而言可能不成立，这是因为关税壁垒和商品差别税率存在，另外一个原因是利润率取决于竞争程度。后者支持“市场定价”(PTM)假说(Krugman(1978))，一些商品在国外市场上出售时的定价等于当地价格，商品定价不会受到汇率影响。Knetter(1989,1993)发现，德国和日本的公司就是一个特别好的印证(例如，德国啤酒出口到英国的价格和出口到美国的价格截然不同)。持久的价格差异可能是因为公司更改价格时有菜单成本，或者是因为在转移产品需求时消费者面临固定成本。这里也存在对类似商品进行分类的问题。即使是麦当劳“巨无霸汉堡包”的消费也取决于具体的条件，如餐馆的位置、排队时间、番茄酱是否免费、汉堡如何与其他商品(如牛奶、可乐等)进行捆绑销售等。有些商品(如汽车、电器)具有不同的国家标准(如110伏特和250伏特)、不同的信息收集成本、不同的关税壁垒政策、不同的环境标准和不同的保证与售后服务。例如，一笔交易的保险费和运输成本估计是交易总额的10%，但比例在不同国家之间存在很大的差异。总

之，当厂商具有差别定价能力时，出口商可能会在出口国进行“市场定价”。

商品套利机会似乎不存在，因为绝对价格在不同国家不相等。Engel 和 Rogers（1996）使用美国和加拿大 23 个城市 14 种不同的价格指数进行了研究，结果表明：

（ⅰ）在同一国家内，同种商品在两个城市之间的相对价格是距离的函数，两者是一种正相关关系。

（ⅱ）当我们对同种商品在不同国家之间进行比较时，价格差异比（ⅰ）中的价格差异大得多；价格差异的“国界效应”等价于国内城市之间 2500 到 23000 英里的距离效应。

（ⅲ）跨国价差比国内价差更具有持久性。

Parsely 和 Wei（1996）及 Rogers 和 Jenkins（1995）得出了与 Engel 和 Rogers（1996）相似的结论，他们也发现跨国价差的持续性远远高于国内价差。Engel（1993）认为绝对 PPP 不存在。他注意到，在 2000 次比较分析中，美国和加拿大相似商品相对价格的（条件）方差比这两个国家内不同商品相对价格（条件）的方差有更大的波动性。对高度同质产品（如出口到美国和英国的螺丝、螺帽和螺栓、纸和玻璃、德国啤酒）的研究也表明绝对 PPP 不成立。

价格指数

绝对 PPP 通常用价格指数来进行检验，但检验在这里存在一些困难。每个国家的指数所覆盖的商品不同，指数中的商品也经常发生改变，因为有些新的商品会被添加到指数之中。在使用时间序列数据时，我们也必须假设绝对 PPP 在基年中成立。严格地说，只有可交易商品才应被包含在 PPP 指数中（除非有理论将可交易商品和不可交易商品的价格联系起来，否则所构造的指数无法像 CPI 指数那样使用）。

绝对 PPP 的一种检验方法是对即期汇率和相对价格进行协整检验：

$$s_t = \alpha p_t - \alpha^* p_t^* + \varepsilon_t$$

对上述方程 OLS 残差的 ADF 检验可以用来对无协整关系的零假设（$\gamma \geqslant 0$）进行检验，

$$\Delta\varepsilon_t = \hat{\gamma}\varepsilon_{t-1} + \sum_{i=1}^{m} \sigma_i \Delta\hat{\varepsilon}_{t-i} + u_t$$

备选择是 $\gamma < 0$。（上述方程可能包含一个常数项（γ 的临界值），这取决于方程的精确形式。）如果 OLS 标准差存在一定的偏差，那么我们无法在单方程框架下对{1，−1}的参数限制进行检验。对这些限制条件的检验要用到 Johansen（1988）的三变量 Var 系统，或使用 Philips 和 Hansen（1990）修正后的 OLS 估计量。一般而言，这些检验方法对协整向量的存在性和间战时期及之前时期的均值反转效应而言是很有利的（对工业国家而言），但对 1973 年后的浮动汇率时期而言就不是很有利（参见 Sarno 和 Taylor（2002）的综述）。

随机游走

单变量单位根检验的问题在于拒绝错误零假设的功效很低（如果根（root）接近于 0.95，那么问题就变得尤其严重）。例如，如果是在根 $(1-\gamma)=0.95$ 时使用 MCS，那么即使观测值个数为 $T=100$，Dickey – Fullert 检验 $\hat{\tau} = \hat{\gamma}/s.e.(\hat{\gamma})$ 也只为 12% 左右。这意味着，即使变量是平稳过程，无法拒绝单位根的概率也约为 88%。因此，很难看出为什么早期的 PPP 检验要使用价格指数的标准单变量单位根检验方法，1973 年后的数据很难拒绝实际汇率的随机游走零假设（除那些通货膨胀率很高的国家外）。

数据扩展可以增加检验功效（尽管增加观察频率（例如，在给定的时期内，将观测频率由年改为月（Shiller 和 Perron（1985）））无法增加检验功效）。因此，研究应对数据集合进行扩展。使用上世纪的数据和单位根检验方法所得到的一致结论是：对通货膨胀适中的国家而言，实际汇率具有长期均值反转现象（Cheung 和 Lai（1994），Lothian 和 Taylor（1996）），即使我们对长期数据中的结构性突变和不同汇率制度（如固定汇率制度、浮动汇率制度和管理浮动制度）的可能性进行考虑后，结论依然成立。

然而，上述研究发现实际汇率也具有较强的持续性，每年约为 15%，这意味着半衰期（half-life）是 3 ~ 5 年（也就是说，对实际汇率的一次冲击而言，要花 3 ~ 5 年时间才能在长期最终价值方向移动一半的距离）。文献综述参见 Rogoff（1996）。近期研究表明，即使是在二战以后，面板数据分析也发现了实际汇率的均值反转特征（即使用 100 多个国家的数据来检验 PPP），半衰期同样约为 4 年（Frankel 和 Rose（1995））。面板数据增加了单位根检验的功效，但仍然可能出现的问题是，面板中的一些国家没有均值反转现象，因为检验结果很容易受“长期内 PPP 成立的一些国家”所主导。当然，相反的情形也是可能的，也就是说，我们无法拒绝非平稳性的零假设，可能的原因是，只有实际汇率是非平稳序列，而其他变量都是平稳序列。Taylor 和 Sarno（1998）及 Sarno 和 Taylor（1998）对这些问题进行了强调，他们发现，与美元不同的是，G5 的实际汇率在 1973 年后的浮动汇率制度时期呈现均值反转特征（使用 Johansen 程序）。尽管标准单位根检验还远不是绝对可靠，但对 20 世纪 90 年代中期而言结论是：实际汇率具有很强的持续性，可能具有均值反转特征（参见 Ardeni 和 Lubian（1991），Grilli 和 Kaminsky（1991），Rogoff（1996））。

我们来看看通过工资价格螺旋用一种替代方法来对 PPP 进行检验的实证结果。我们在附录里分析到，国外（进口）价格上升或国内货币贬值使国内产业的生产成本上升，从而使国内价格和工资上涨。解释国内通货膨胀的方程是：

$$\Delta p=\frac{b_1}{1-b_1}[f+b_1(\chi_w-\chi_p)+a_2(y-\bar{y})]+\Delta(p+s) \tag{5}$$

其中，$(y-\bar{y})$ 是产出相对于自然生产率的偏差，f 是“工资推动”因子，χ_w 是外生的实际工资增长率，χ_p 是劳动生产率的增长率。如果

$$f+b_1(\chi_w-\chi_p)+a_2(y-\bar{y})=0 \tag{6}$$

那么 PPP 成立。因此，当产出按照自然增长率进行增长、“工资推动”因子 f 等于 0、实际工资增长等于劳力生产率时，PPP 成立。我们可以看到，（6）中因子涉及长期缓慢变动的复杂经济、社会和政治力量，这可以说明（即使使用长期数据）实证 PPP 存在的困难。

一些商品（如汽车和电器）的市场定价（pricing to market，PTM）也会导致 PPP 产生偏离（Krugman（1978））。如果一些厂商的出口商品用外币计价，那么汇率的任何变动在开始时候都不会对相对价格产生影响，因此，LOOP 被违背，实际汇率和名义汇率一起变动。然而，尽管 PTM 会发生在一些商品上，但它似乎并没有得到广泛采用（美国除外）（Goldberg 和 Knetter（1997））。

单位根和非线性

用多种方法研究的结果都表明，实际汇率没有均值反转特征。Rogoff（1996）注意到，使用 Harrod-Samuelson-Balassa（HSB）效应预测结果是 PPP 并不总是成立，因为经济快速增长的国家理应使实际汇率提高。在最简单的形式中，HSB 效应假设不可交易部门的生长率比可交易部门的生产率增长得更快。LOOP 在可交易部门中成立，但生产率的快速增长使得可交易部门的工资出现上涨（可交易商品的价格保持不变，因为单位成本没有变化）。竞争性劳动力市场使得不可交易部门的工资上涨，因为生产率没有变化，所以不可交易商品的价格上涨。后者使得价格指数（如 CPI）上升，如果汇率没有变化（因为可交易商品的价格不变），那么总量价格的使用使得 PPP 被违背，国内实际汇率上升。这很符合像日本这样的国家和日元 - 美元汇率（1945—1996）的实际情形，但这种解释对其他工业化国家不合适，HSB 效应的证据不尽一致（如 deGregorio、Giovannini、Wolf、Gordon 和 deMenil（1994））。

其他学者（如 Krugman（1990））认为累积的经常账户赤字应对实际汇率产生影响，因为它们代表的是资产组合平衡效应和财富在不同国家之间的重新分配。另一个分支的文献（Froot 和 Rogoff（1991））则认为，政府消费的主要是不可交易商品，这会导致可交易商品的相对价格发生变化，从而使得实际汇率上升。

考虑交易成本后的 PPP 模型（Dumas（1992），Senu、Uppal 和 vanHulle（1995））则强调，在一个无套利的“交易成本区间”内，实际汇率可能服从随机游走，但在区间之外，潜在的套利利润是如

此之高以至于均值反转行为会发生。总之，实际汇率服从一个带均值反转的非线性过程。时间加总使得个体商品在不同状态之间的不连续调整被匀滑掉，如果我们将那些套利成本不同的商品进行加总，那么用于检验 PPP 的总量价格指数似乎服从匀滑的非线性过程。

另一种观点（Kilian 和 Taylor(2003)）认为，非线性的产生是因为人们对“准确的”汇率水平有着不同的预期。因此，当汇率接近均衡汇率、各种观点四处扩展时，基本面交易者会受噪声交易者（如相信图表和技术分析的投资者）的影响，这会产生汇率的单位根行为。当汇率与均衡值有很大的偏离时，更多的基本面交易者会意识到这种偏离，从而使汇率重新回到均衡状态（也可以参见 DeGrauwe、Dewachter 和 Embrechts(1993)的类似模型，模型给出的是汇率非线性的“混沌”行为）。

然而，尽管基于标准单位根检验的长期 PPP 作为一个经验事实已广为接受，PPP 之谜仍然存在。也就是说，实际汇率和名义汇率在短期内的波动率相同，这表明所有“行动”都是相对于名义汇率而不是相对价格而言的。这意味着，相对于实际冲击（如对生产率和偏好的冲击，人们认为这些变量的波动性不是特别的高）而言，货币冲击和金融冲击（即名义冲击）在这里要更为重要一些。但是，因为粘性工资和价格会进行调整，所以我们期望这些名义冲击的影响在两三年后完全消失——事实上，与 PPP 实际偏差的存在时间长于五年。这就是那个谜，也就是说，是什么原因导致了实际汇率出现剧烈的长期波动呢？

走出上述困境的一种途径是，假设汇率服从一个带均值反转的非线性过程，使用蒙特卡罗试验来生成模拟数据，然后对传统单位根检验的功效（也就是说，拒绝错误零假设的能力）进行考察。Taylor(2001)指出，在这种条件下，半衰期的估计存在很大的向上偏离。当真实模型是 AR(1)但在进行单位根检验时所使用的数据数却是时间加总后的数据时（例如，真实 AR(1)模型使用的是月度数据，但实际数据却是年度数据），半衰期也会大幅增加。

近期的实证工作使用实际汇率的非线性模型（如 Obstfeld 和 Taylor(1997)，Michael、Nobay 和 Peel(1997)，针对间战时期）。例如，Taylor、Peel 和 Sarno(2001)对 1973 年后与美元相关的四对双边汇率使用了 ESTAR 模型（详见第4章）。ESTAR 模型（假设服从 AR(1)过程）的一个简化版本如下：

$$z_t = (\pi_0 + \pi_1 z_{t-1}) + (\pi_1^* z_{t-1}) F(z_{t-j}) + \varepsilon_t$$

其中，$F(z_{t-j}) = 1 - \exp[-\gamma(z_{t-j} - c)^2/\sigma_z^2]$。$F(z_{t-1})$具有 U 形形状，在 0 和 1 之间取值。如果 $z_{t-j} = c$，那么 $F = 0$，与 z_t 有关的上述方程变成一个线性 AR(1)模型。当$(z_{t-j} - c) \to \infty$时，不均衡的程度很高且 $F = 1$，因此我们就得到了一个不同的 AR(1)模型，z_{t-1}的系数为$(\pi_1 + \pi_1^*)$。模型容许 $\pi \geqslant 0$，但我们必须有$(\pi_1 + \pi_1^*) < 0$。对很小的偏差而言，z_t 服从单位根过程（或更为发散的过程），但对大的偏差而言，这个过程是带有均值反转的过程。如果 ESTAR 模型是正确的模型，那么我们马上就可以看出为什么传统的（线性）单变量单位根检验会导致一些错误的推断。一个有点自相矛盾的结果是，当 ESTAR 模型是正确的模型时，无法拒绝单位根则意味着实际汇率与均衡值很接近。并不令人惊讶的是，MCS 结果表明，当 ESTAR 模型是正确的选择模型时，单变量检验的功效很低，但面板单位根检验确实能提高校验功效。

在非线性模型中，实际汇率冲击的半衰期取决于最初的非均衡位置和冲击大小。Taylor、Peel 和 Sarno(2001)使用 MCS 的研究结果表明，当冲击较小时，*ESTAR* 模型所给出的半衰期约为 3 到 5 年，但对最大的冲击而言，半衰期小于 3 年。这为“冰河”（Rogoff(1996)）半衰期之谜提供了解决方法，半衰期之谜是在 PPP 的线性调整模型中发现的。

在过去的年份中，PPP 检验的“技术能力”与日俱增，研究结果毫无疑问为均值反转和长期 PPP 提供了有利证据。然而，即使存在指数问题和单位根检验的低功效问题，在出现上述最新进展之前，绝大多数经济学家和政策制定者本应该持实际汇率可能是均值反转的观点。政策问题始终是什么原因导致了实际汇率出现持续大幅上升，实际汇率的上升意味着可交易部门所在的产业会遭到破坏，即使这些产业的生产率和单位成本（以国内货币计价）都保持不变（与之相伴的事实是，当实际汇率下降且市场份额不能完全重新获得时，这种情形可能不是对称的）。肯定

为真的是国际商品市场的整合程度比国内商品市场低得多（这是大家经常引用的加入共同货币区域（如欧洲货币联盟）的原因之一，由于竞争加剧，共同货币区域的“价格透明性”鼓励了国内贸易，价格不一致的程度也要小得多）。

24.3 抵补利率平价

外汇市场上主要有两类“交易”。第一种交易涉及“即期”汇率，“即期”汇率是购买者进行直接交割的汇率报价（交割事实上是在两个工作日后进行）。第二种交易涉及远期汇率，远期汇率是购买者今天承诺的在将来某个时期进行交割的价格。你可以通过今天的远期市场来“锁定”未来某个日期的汇率，从而为外币支付票据进行套期保值。对绝大多数主要货币而言，成交频率最高的是期限为 1 到 6 个月的远期合约，在极为特殊的情形中，3 到 5 年的远期合约成交频率也很高。外汇市场上的做市商主要是大银行，他们都在场外市场（OTC）上进行交易。

即期汇率和远期汇率之间的关系描述如下。假设英国一家公司的财务主管拥有 A 英镑的现金，他可以在英国或美国进行为期 1 年的投资，所有收益都必须支付给公司股东。我们假设远期交易没有任何违约风险。因此，对于财务主管而言，他并不关心在哪个地方进行投资，从而在英国的投资收益应等于在美国的投资收益。他在英国的投资收益为 $A(1+r)$，r 是英国的利率。在美国的投资用英镑计算的收益可以通过即期汇率 S（英镑/美元）和一年前的远期汇率 F 来进行计算。将 A 英镑兑换成美元后就得到 A/S 美元，如果用 r^* 表示美国的利率，那么一年后这笔资金会增值到 $(A/S)(1+r^*)$ 美元。如果进行交割的一年期远期汇率为 F（英镑/美元），那么英国的财务主管在今天就可以将汇率“锁定”，在一年后可以确定地收到 $(A/S)(1+r^*)F$ 英镑。使收益相等就得到

$$A(1+r)=(A/S)(1+r^*)F \tag{7}$$

上式可以简化为

$$F/S=(1+r)/(1+r^*) \tag{8}$$

或

$$f-s=r-r^* \tag{9}$$

其中，$f=\ln F$，$s=\ln S$，我们在上面的计算中使用了近似等式 $\ln(1+r)=r$，这里的用小数进行表示。上述方程表示的是“抵补利率平价”（CIP）条件，它是基于无风险套利的均衡条件。如果 CIP 不成立，那么就存在力量使得均衡得以重新恢复。例如，如果 $r>r^*$ 且 $f=s$，那么无风险套利利润存在。美国居民在今天可能购买英国短期国债，推动债券价格上涨，从而使利率下跌。当然，美国居民也可以在今天购买即期英镑，卖出远期美元，从而即期英镑升值（即 s 下跌），f 上升，因此方程（9）倾向于重新成立。事实上，因为交易无风险，所以套利交易者的行为使得远期汇率报价等于 $s+r-r^*$。

若使用的是连续复利则会有 $F=Se^{(r-r^*)T}$，其中 T = 远期合约的到期时间（用年表示）。从而（9）就变成了一个精确的（而不是逼近）。

24.4 非抵补利率平价

我们来重复上一节的分析，但这里假设那位英国的财务主管愿意对一年后的汇率 S^e_{t+1} 进行预测，她希望在一年后将美元投资兑换为英镑。如果财务主管属于风险中性类型，那么她只关心两种投资的预期收益。因此，她会继续在美国进行投资而不是在英国进行投资，直到美国和英国投资的预期收益相等为止。

$$S^e_{t+1}/S_t=(1+r_t)/(1+r^*_t) \tag{10}$$

$$s^e_{t+1} - s_t = r_t - r_t^* \tag{11}$$

其中，$s_{t+1} = \ln S_{t+1}$。（如果使用连续复利利率，那么方程（11）就是一个精确等式。）非抵补利率平价（UIP）可以解释为风险中性假设下资本账户的一个均衡条件，因为如果UIP成立，那么在两个国家之间转移投机资金不会有任何好处。英国的财务主管知道她承担了风险，因为一年后的汇率是不确定的，然而，在进行资产组合配置决策时，她可以忽略这种风险。

现在，我们通过CAPM模型来放松风险中性假设。对英国的财务主管而言，无风险利率为r，在美国进行风险投资的预期收益为

$$1 + E_t R_{t+1}(UK \rightarrow US) \equiv S^e_{t+1}(1 + r_t^*)/S_t \tag{12}$$

根据CAPM模型，我们有

$$E_t R_{t+1} - r_t = \beta_i (E_t R_{m,t+1} - r_t) \tag{13}$$

其中，β_i 是国外投资的 β 值（取决于市场投资组合和美国投资组合之间的协方差）。$E_t R_{m,t+1} - r_t$ 是市场投资组合的预期收益，市场投资组合包括所有的货币和资产（也就是说，世界资产组合）。（13）的右边是由CAPM模型所给出风险溢价度量指标。目前要注意的是，如果我们假设 $\beta_i = 0$，那么（13）就简化为UIP。CAPM模型是SDF模型的一种特殊情形，它适用于国内和国外的资产收益（参见第13章）。

再回到UIP方程（11），显然，如果该方程不成立，那么风险中性投资者在不同国家之间进行资金转移就能获得收益。如果资金转移是在很短的时间内发生，那么UIP在任何时点都成立。显然，UIP条件假设市场由风险中性投机者所主导，风险规避型“理性投机者”和噪声交易者对市场价格都没有太大的影响。

24.5 远期汇率无偏性

由（9）和（11）可知，如果CIP和UIP同时成立，那么远期汇率是对未来即期利率的无偏预测，

$$f_t = E_t s_{t+1} \tag{14}$$

注意，无论 Es_{t+1} 的形式如何，无偏性都成立（即我们不需要理性预期假设），但无偏性确实需要风险中性假设（目的是使UIP成立）。如果UIP、CIP和远期汇率无偏性（FRU）中有两个成立，那么另外一个也成立。

在风险中性假设下，如果FRU不成立，那么远期市场的投机活动存在（风险）获利机会。无论（14）成立是因为远期市场上存在活跃的投机活动，还是因为CIP成立且所有投机活动都发生在即期市场上（从而UIP成立），结果都与EMH无关。关键是不存在获利机会。

24.6 实际利率平价

如果UIP和PPP都成立，那么

$$\Delta s^e_{t+1} = s^e_{t+1} - s_t = (r - r^*)_t \tag{15}$$

$$\Delta s^e_{t+1} = \Delta p^e_{t+1} - \Delta p^{*e}_{t+1} \tag{16}$$

于是，

$$r - \Delta p^e_{t+1} = r_t^* - \Delta p^{*e}_{t+1} \tag{17}$$

因此，

$$\text{PPP} + \text{UIP} \Rightarrow \text{实际利率平价(RIP)}$$

如果UIP、PPP和实际利率平价（RIP）中有两个成立，那么另外一个也成立。如果商品市场的信息成本和调整成本相对较高，那么不足为奇的是，PPP只是在相对较长的时期中（比如说5～10

年）才成立。事实上，我们知道实际汇率在短期内存在大幅变动。因此，即使在风险中性假设下（即 UIP 成立）人们也会认为：预期的实际利率平价只在相当长的时期中才成立。注意，相等的是预期实际利率。然而，如果我们假设代理人在数年中进行通货膨胀和汇率变动预测时不会犯系统性错误，那么实际平均利率相等。

RIP 条件在所谓的国际 Fisher 假说下也成立，它可以视为一种套利关系，“资本”（即投资资金）在国家之间的流动使得每个国家的预期实际收益相等。假说认为，每个国家具有代表性的一篮子商品（价格为 p 和 p*）为国际投资者带来的效用都相同（例如，来自英国的哈罗斯 Hamper 与来自纽约的萨克斯 Hamper 可以视为相同）。国际投资者通过购买金融资产或直接投资的方式将资金转移到那些实际收益最高的地方。注意，投资收益的增长是用某个特定国家的消费品来进行表示的。因此，如果实际收益是在英国获得的，（比方说）但你希望消费美国商品（如在美国生产的萨克斯 Hamper），那么你必须在投资期末将英镑兑换成美元。然而，在 RIP 条件下价格竞争力成立，所以你在美国和英国有着相同的购买力。

值得强调的是，上面讨论的所有关系都是均衡条件，但因果关系的方向不是很明显。因此，在 UIP 情形中，我们不能说是利息之差“导致”了汇率的预期变化（反之亦然）。当然，我们可以对模型进行扩展，从而将一些因果关系假设嵌入到模型之中。例如，假如我们断言（基于经济学原理和政府行为证据）中央银行货币供给量的外生性变动会“导致”国内利率发生变化。那么，在给定的 UIP 条件下，货币供给量也会“导致”汇率预期升值或贬值速度发生变化。货币供给量的外生性变动对国内利率和汇率的预期变化都会产生影响。“货币”在这里是原因（根据假设），UIP 关系中的变量由模型同时决定。

根据原理，当我们对 UIP、PPP 和 FRU 三种关系或 UIP、PPP 和 RIP 三个条件的有效性进行检验时，我们只需对其中的两个进行检验。因为如果其中两个成立，那么另外一个也会成立。然而，由于数据的可获得性和各个变量的数据质量差异，我们要同时对三种关系进行检验。

24.7 小结

* 如果 PPP 成立，那么价格竞争力和实际汇率都是不随时间变动的常数。无论是在国内还是在国外购买商品，商品的成本都相同。

* 有证据表明，实际汇率具有长期波动特征，这意味着价格竞争力在 5 ~ 10 年后会恢复到正常水平。

* 如果远期汇率由 $F_t = S_t(1 + r_t)/(1 + r_t^*)$ 给出，那么无风险套利利润就不可能产生。这就是 CIP 条件。

* UIP 意味着国内货币的预期贬值率等于国内外的利息差异。

* CIP 和 UIP 意味着远期汇率是未来即期汇率的无偏预测。

* UIP 和 PPP 意味着不同国家的实际利率都相等。

附录：PPP 与工资 - 价格螺旋

在附加工资预期的菲利普斯曲线中，工资增长率 Δw 由物价上涨率 Δp 和超额需求 $(y - \bar{y})$ 来决定。在此基础上，我们可以考虑如下可能性：工人在观察到生产率后会推动实际工资的增长 χ_w。其他因素 f（如规定最低工资的法规、社会经济力量）也会对工资产生影响。人们通常假设价格由单位工资成本和国内的原材料进口价格来决定。因此，我们的工资 - 价格模型为

$$\Delta w = \chi_w + a_1 \Delta p + a_2(y - \bar{y}) + f \tag{A1}$$

$$\Delta p = b_1(\Delta w - \chi_p) + b_2 \Delta pm \tag{A2}$$

其中，χ_p 是劳动生产率的趋势增长率。假设进口主要是同质的可交易商品（如农产品、石油、铁矿石、煤）或资本品。这些进口商品的国外价格由世界市场决定，可以按照如下关系转换成国内价格：

$$\Delta pm = \Delta p^{*} + \Delta s \tag{A3}$$

将（A1）代入到（A2）后我们就得到

$$\Delta p = (1 - a_1 b_1)^{-1}[b_1(\chi_w - \chi_p) + b_2 \Delta pm + a_2 b_1(y - \bar{y}) + b_1 f] \tag{A4}$$

方程（A4）是附加价格预期的菲利普斯曲线（PEAPC），它将价格增长率与超额需求（$y - \bar{y}$）和其他变量联系起来。如果我们进行合理假设：货币幻觉在长期中不存在（$a_1 = 1$），也就是说，长期 PEAPC 是一条垂线，且总成本具有齐次性（$b_1 + b_2 = 1$），那么（A4）就变成

$$\Delta p = [b_1/(1 - b_1)][f + b_1(\chi_w - \chi_p) + a_2(y - \bar{y})] + \Delta(p^{*} + s) \tag{A5}$$

假设中括号中的项都等于零，那么 p^{*} 和 s 会对国内价格产生长期影响，从而 PPP 成立，也就是说，

$$\Delta p = \Delta p^{*} + \Delta s \tag{A6}$$

国外价格 p^{*} 的上涨或国内货币的贬值（即上升）都会导致国内价格出现上涨（通过（A2）），这反过来又导致更高的工资增长率（通过（A1））。因为工资的上涨会进一步导致价格上涨，所以工资－价格反馈作用的强度取决于 a_1 和 b_1 的大小。在齐次性假设下，$a_1 = 1$，$b_1 + b_2 = 1$，反馈作用使得 PPP 在长期中成立。也就是说，在其他条件均保持不变的情况下，国内货币贬值（或国外价格上升）1% 最终会导致国内价格指数上涨 1%。当然，在这个模型中，PPP 在短期内通常不成立，原因要么是货币幻觉的存在（$a_1 < 1$），要么是总成本 $b_1 + b_2 < 1$，也可能是（A5）中中括号内各项的影响所致。

第25章

检验 CIP、UIP 和 FRU

目的

∗ 说明抵补利率平价(CIP)检验如何注意数据问题,因为抵补套利(实际上)是无风险套利。

∗ 使用单方程和 Var 框架来对远期汇率的无偏性(FRU)和非抵补利率平价(UIP)进行检验。

∗ 说明“比索问题”如何导致 FRU 检验的“解释”变得非常复杂。

在这一章里,我们讨论抵补利率平价、非抵补利率平价和远期汇率无偏性的检验方法。我们发现,就所研究的绝大多数期限和时期而言,抵补利率平价获得了很强的支持,而远期汇率无偏性(FRU)(或非抵补利率平价 UIP)的检验结果与假设相反,我们使用调查数据来摸索一些检验方法,目的是为了确定上述结果是否是因为风险中性或 RE(理性预期)的不成立。在这一章里,我们讨论 FRU 在单期上中的单方程检验和使用 Var 框架的多期检验。

25.1 抵补套利

我们来讨论通过抵补套利来获取无风险利润在实践中是否可能的问题。在现实世界中,对即期利率、远期利率和即期汇率而言,买卖价差在评估潜在的无风险获利机会时是极其重要的。在最严格的定义中,套利交易要求没有资本金:投资者借入资金。考虑一位英国投资者,他以 $r^0_{英镑}$ 的要价在欧洲英镑市场借入 A 英镑,期末欠款总额为

$$Z_1 = A\left(1 + r^0_{英镑}\frac{D}{365}\right) \tag{1}$$

其中,A 是借入金额(英镑),Z_1 是期末欠款金额(英镑),$r^0_{英镑}$ 是欧洲英镑贷款的报价(用比例表示),D 是资金的借贷天数。

现在考虑如下系列交易。投资者在即期市场上以出价汇率 S^b(美元/英镑)将 A 英镑兑换成美元,因此,S^b 是即期市场上英镑的买入价格。这位投资者将这些美元以 $r^0_{美元}$ 的出让利率投资于欧洲美元存款。与此同时,他又以远期汇率 F^0(基于报价方)将这些美元转换成英镑。所有这些交易都在瞬间内发生。D 天后他能确定地收到的英镑数为

$$Z_2(\text{UK}\rightarrow\text{US}) = \frac{A\cdot S^b[1 + r^b_{美元}(D/360)]}{F^0} \tag{2}$$

注意美国的天数计算惯例,(2)中将“1 年”定义为 360 天。将 A 英镑投资于美国资产然后再在远期市场上将美元兑换成英镑所获得的百分比超额收益 ER 为

$$\begin{aligned}\text{ER}(\text{英镑}\rightarrow\text{美元}) &= 100\left(\frac{Z_2 - Z_1}{A}\right)\\ &= 100\left[\frac{S^b}{F^0}\left(1 + r^b_{\text{美元}}\frac{D}{360}\right) - \left(1 + r^0_{\text{英镑}}\frac{D}{365}\right)\right] \qquad (3)\end{aligned}$$

因此，如果你将 M 英镑兑换成美元来进行套利，那么你在期末可以确定获得的英镑利润为（$M \cdot$ ER/100）英镑。再从一位美国居民的角度来看抵补套利交易，我们可以考虑如下交易的抵补套利收益：先以即期汇率将美元兑换成英镑，然后在英国进行投资，最后以当前的远期汇率将英镑兑换成美元。我们必须将所得收益与美国美元资产的借贷成本进行比较。与(3)类似的计算公式为：

$$\text{ER}(\text{英镑}\rightarrow\text{美元}) = 100\left[\frac{F^b}{S^0}\left(1 + r^b_{\text{英镑}}\frac{D}{365}\right) - \left(1 + r^0_{\text{美元}}\frac{D}{360}\right)\right] \qquad (4)$$

假设套利是无风险套利，那么我们期望 ER（英镑→美元）和 *ER*（美元→英镑）都等于零。抵补套利不涉及“价格风险”，唯一的风险就是信用风险，因为交易对手有可能无法支付利息或进行远期货币交割。如果我们要对 CIP 假说进行检验，那么我们需要同时获得即期汇率和远期汇率“交易”的报价和两个国家的利率。由于存在抵补套利，所以大量研究试图找出可能的获利机会，但不是所有研究都使用即时交易汇率数据。然而，Taylor（1987，1989*a*）对外汇市场上“平静”时期和“动荡”时期中的 CIP 关系进行了考察，他使用的是外汇市场和货币市场上经纪人的即时报价数据。因此，我们集中讨论他的研究。Taylor 所使用的汇率代表公司进行买卖的报价，因此，也代表任何时点上在市场中所能获得的最佳价格（最高卖价，最低买价）。与之相反的是，路透电子屏幕上的报价通常是为了便于传递信息，可能不是实际成交价格。Taylor 使用的是欧洲货币汇率，交易对手几乎没有信用风险，因此，唯一的差别就是计价货币的名称不同。

Taylor 还考虑了经纪人的佣金，在如下假设下对上述收益重新进行了计算：欧洲货币交易的佣金大约是1%的1/50。例如，如果考虑佣金，那么借入欧洲美元的利息成本为 $r^0_{\text{美元}} + 1/50$，而所有欧洲美元存款的利息收益都会减少一个相同的数量，即 $r^b_{\text{美元}} - 1/50$。

Taylor 认为即期交易和远期交易的佣金费用是如此的低以至于可以忽略不计。在1987年的研究中，Taylor 使用的是1985年11月11日、12日、13日每隔10分钟的数据。这些数据表明，3500个潜在的套利机会存在，在考虑佣金成本后，他发现抵补套利机会几乎没有获利。因此，结果为抵补利率平价和有效市场假说提供了强有力的证据。在后来的研究中，Taylor（1989a）对抵补套利关系进行了重新考察，但这次考察的是外汇市场的“动荡”时期。历史时期选择为1967年11月的英镑贬值时期、1972年6月的英镑浮动时期及20世纪80年代英国和美国的普选时期。抵补套利收益计算的期限分别为1个月、2个月、3个月、6个月和12个月。结论要点如下：

* “动荡”时期存在一些获利机会。

* 与20世纪60年代的固定汇率时期相比，浮动汇率时期的获利空间更小，1972年以后，当参与者对浮动汇率有一些经验之后，获利空间进一步缩小。

* 连续几个时期中套利获利机会的频率、规模和持续性都随合约期限的增加而增加。因此，与1个月的远期交易相比，12个月远期交易的获利要更多一些，获利频率也更高一些。

我们来考虑一个特殊例子。1967年11月，将1百万英镑兑换成美元可以获利473英镑，但在英镑贬值后（即“动荡”时期之后），3个月和6个月期限的无风险套利分别能产生4000和8000英镑的收益。20世纪60年代生效的资本管制（针对英镑外流）也无法解释这些结果，因为欧洲英镑存款/贷款并不受其管制。在动荡时期，期限较长合约的收益能持续一些日子，但期限较短合约的利润要小得多。

一月期合约小而持续的收益可能是因为交易者的机会成本为正。市场上认为值得花费时间和精力去利用非常小的获利机会的交易者不是很多。在特殊市场上的投入时间受到限制的条件下，交易者可能偏好有较大预期收益的交易，即使这种交易伴随着风险（如在未来即期市场上进行投机），交易者的偏好还是这样。对他们而言，更值得他们做的是聚集在一起进行相互交

流，而不是使用获利微小的机会。在更长的时期中，交易者在市场上可获得的无风险收益相当的高，这明显是对有效市场假说的违背。为什么会出现这种情形呢，Taylor 提出了好几种假说，所有假说基本上都是基于交易者在外汇市场上的信用头寸限制。

做市商通常不是无条件地与交易对手达成任意数量的交易。银行的管理层一般会对交易另一方的银行进行限制（即考虑信用风险），银行对对方持谨慎态度，对对方的最大负债规模进行限制。从而，抵补套利存在流动性约束。一旦信用限额“满”了，与这家银行的业务物往来也就中止了（直到所有未清债务全部清偿为止）。因为资金“逃逸”相对频繁，所以短期市场表现出抵补套利偏好。

银行也不愿意让其外汇交易商向其他银行进行长期限（如 1 年）大额外币借款。例如考虑英国的一家外汇交易商，他从纽约一家银行借入一大笔美元来进行一年期限抵补套利交易。如果英国银行也想从这家纽约银行为其客户借入美元，那么英国银行会因为信用限额而使得这项工作很难开展。如果是这样，那么外汇交易商会与其他银行达成一个松散的信用限额，但这会对长期限的抵补套利形成限制。交易商进行自我信用限额的另外一个原因是中央银行经常要求商业银行定期公布财务报表，在对商业银行的“平稳性”评估中，中央银行可能会考虑商业银行的短期头寸状况。如果外汇交易商借入大笔资金来进行抵补套利，那么这会在短期头寸中有所体现。

Taylor 注意到，只要本金超过 1 亿美元，一些大银行愿意支付比欧洲美元存款利率高 0.0625%（1% 的 1/16）的利率。因此，Taylor 意识到，他所使用的欧洲美元利率可能存在度量误差，获利机会可能比他研究中所发现的机会多或少。

表 1 抵补套利：百分比超额收益（1987 年）

	1月		6月		1年	
	（英镑→美元）	（美元→英镑）	（英镑→美元）	（美元→英镑）	（英镑→美元）	（美元→英镑）
星期一 1987.6.8 中午 12 点	-0.043	-0.016	-0.097	-0.035	-0.117	-0.162
星期二 1987.6.9 中午 12 点	-0.075	-0.064	-0.247	+0.032	-0.192	0.150

数据来源：Taylor（1989a）数据摘要，表 3。

Taylor 发现，20 世纪 60 年代固定汇率时期的抵补套利收益相对而言要高一些；然而，浮动汇率时期的抵补套利收益要小得多，出现频率也低得多。例如，从表 1 我们可以看出，1987 年就一个月的期限而言，将英镑兑换成美元几乎没有获利机会。然而，就一年的期限而言，在星期一和星期二将美元兑换成英镑都能获得无风险套利利润，1 百万美元在这里的利润为 1500 美元。

Taylor 的研究没有考虑国内外投资利息收益的税收差别，这可以解释一年期抵补套利持久利润的存在。不大可能出现的是，市场参与者受（比方说）欧洲英镑和欧洲美元投资之间相对违约风险的影响，因此这不可能解释一年期的套利利润。注意，对具有不同信用风险特征（要么是“市场价格风险”，要么是信用风险）的资产而言，我们不可能充分地检验他们之间的 CIP 关系。例如，比较欧洲英镑存款和美国公司债券之间抵补交易的研究不可能为远期市场效率和 CIP 提供大量信息。

注意，即使交易事实上没有发生，CIP 也可能成立。这对保证 CIP 成立的无风险套利来说是一个很大的威胁。这在如下事实中可以部分地得到反映：如果你去一家银行询问远期报价，那么

银行提供给你的是用CIP关系所计算的远期汇率。也就是说，银行在查对$r_{英镑}$、$r_{美元}$和S_t的值后会向你报出远期汇率F_t：

$$F_t = S_t\left(\frac{1 + r_{美元}}{1 + r_{英镑}}\right) \tag{5}$$

我们在这里忽略了买卖价差。如果银行所报的远期汇率F_t与CIP方程所给出远期汇率不同，那么银行显然不是理性的或者说是极其愚蠢的。使用做市商实际交易的数据来寻找潜在的获利交易显然是检验CIP的有用方法。然而，有关CIP的早期研究都是进行如下回归：

$$(f - s)_t = a + b(r_{美元} - r_{英镑}) + \varepsilon_t \tag{6}$$

CIP的零假设是H_0: $a = 0$且$b = 1$，如果存在交易成本，那么这可以用$a \neq 0$来体现。因为($r_{美元} - r_{英镑}$)是外生的，所以(6)的估计应使用2SLS或IV而不是使用OLS。然而，CIP的这些回归检验存在很多尖锐的问题。回归一般不能区分买价和卖价，不会明确地（或仔细地）考虑交易成本，所使用的利率序列经常不属于同一个样本期。此外，即使你没有拒绝零假设：$a = 0$且$b = 1$，这也只意味着CIP在平均意义下成立，但这并不意味着它会连续成立。出于这些原因的考虑，我们不报告回归检验的结果。

Clinton(1988)指出，实际结果与CIP之间的偏差不会比两个存款市场或外汇互换市场（即远期交易的实际实施，参见Cuthbertson和Nitzsche(2001b)）的最低成本还要高。他用1985—1986年期间的数据所估计的区间带为每年$+0.06\%$。

Balke和Wohar(1998)及Peel和Taylor(2002)使用非线性TAR模型来考察CIP偏差的动态行为。Peel和Taylor(2002)使用间战时期的周数据所估计的带宽（每年）为50个基点，在带的外部存在适度的持久性。Balke和Wohar(1998)使用1974年1月～1993年9月美英数据发现，在中心带之外具有（渐近）均值反转特征，而带内也存在持久性，但持久性要弱一些。总之，有证据表明，就最近的数据时期而言，CIP偏差所导致的套利利润相对而言似乎要更少一些。

25.2 非抵补利率平价

国内投资者（如"欧洲居民"）以r^*的年利率在国外无风险（美国）债券上投资1英镑后将收益在一年后兑换成国内货币的投资收益为$(1 + r^*)(S_{t+1}/S_t)$，这里的S是用单位外币表示的国内价格（如英镑/美元）。如果是在美国进行投资，那么国内投资者的超额收益（高出国内无风险资产投资收益的部分）为：

$$1 + R_{t+1} = \frac{(1 + r^*)S_{t+1}}{(1 + r)S_t} \tag{7}$$

$$\ln(1 + R_{t+1}) \approx R_{t+1} = (r^* + \Delta s_{t+1}) - r$$

这里使用了$\ln(1 + x) \approx x$和$\ln(S_t) \approx s_t$。国外投资的收益等于国外利率加上外币的预期升值。从事前来看，国外投资的（渐近）期望超额收益为

$$E_t R_{t+1} = (r^* + E_t \Delta s_{t+1}) - r \tag{8}$$

其中，$E_t \Delta s_{t+1} \equiv E_t s_{t+1} - s_t$，$s_{t+1}$是唯一的随机变量。注意，$\Delta s_{t+1} > 0$，也就是说，$s$的增加意味着外币升值（本币贬值）。由于未来的外币收益无法在远期市场上进行"抵补"，所以投资外国资产是有风险的。如果投资者属于风险中性类型且只关心预期收益，那么套利可以保证预期超额收益$E_t R_{t+1} = 0$，因此，我们就得到了UIP（非抵补利率平价）条件：

$$r - r^* = E_t \Delta s_{t+1} \tag{9}$$

UIP意味着，平均而言，你（比如说）在美国和欧洲之间转移资金不可能获利。你可能在一些"赌博"上会赢，但你也可能在一些赌博上输，净收益平均为零。现在考虑一位在"欧元区"进行投资的美国居民。对这位美国投资者而言，国外利率就是欧元利率，令$r^* = 6\%$，美国国内利率$r = 2\%$。于是，如果UIP成立，那么美国投资者在欧元上的额外收益（相对于美国利率而言）将被欧

元相对于美元 4% 的贬值(平均而言)所抵消。换句话说,国外(即欧洲相对于美国)的高利率并不意味着国外投资能获得高收益,它为外币同等预期贬值提供了信号。

当然,对那些高通货膨胀国家而言(例如,20 世纪 80 年代的南美洲国家、20 世纪 90 年代后期的东南亚国家和 20 世纪 90 年代早期的东欧经济体),可以在很宽的范围预测无条件 UIP 关系会出现。利率相对较高的国家也会出现货币贬值。对像英国、德国、瑞士和日本这些通货膨胀适中的国家而言(Engel(1996)),UIP 基本上是无条件地成立。然而,这只是对 UIP 一个非常弱的检验,因为重要的是检验涉及条件期望,也就是说,在某个时点利息差异已知的基础上,资金转移获利是否只是暂时的。一个合适的回归方程是

$$\Delta s_{t+1} = \alpha + \beta(r - r^*)_t + \gamma\Lambda_t + \varepsilon_{t+1} \tag{10}$$

其中,Λ_t 是 t 时刻已知的变量。如果 UIP + RE 成立,那么我们期望 $\alpha = \gamma = 0$、$\beta = 1$ 且 ε_{t+1} 序列不相关。经常发现的结果是 $\beta \neq \pm 1$,事实上 β 一般为负,对不同货币(美元是计价货币)、不同时期和 Δs_{t+1} 的不同期限(如月、季、年)而言,β 的范围为 -1.0 到 -4.0。正如我们在下面即将看到的那样,(10)就是 FRU 方程,因为在 CIP 条件下,我们可以用远期升水 fp_t 来替换 $(r - r^*)_t$。

实证结果表明 $\beta < 0$,这意味着,在高利率差异有利于国外的时期中,外币会倾向于升值,这($\Delta s_{t+1} > 0$)为那些借入美元贷出外币的美国投资者带来正的美元收益。因此,UIP 遭到拒绝。然而,在使用年度数据的回归方程中,R^2 通常都很小(即为 0.035),因此,这样的"赌博"风险极高,尽管在一轮赌博中他们还是有所"收获"(收益为正)。

注意,$\beta < 0$ 的上述回归(10)并不意味着平均来看简单地持有那些利率比美国高的国家(如 20 世纪 90 年代像俄罗斯、印度尼西亚、巴西、阿根廷、土耳其这样的高风险国家)的存款就能赚钱。这是因为,"截距"可能抵消利差项的影响。方程(10)表明,持有利率比美国高的国家的存款能赚钱(即回归方程 $Y = \hat{\alpha} + \hat{\beta}X$ 可以写为 $(Y - \bar{Y}) = \hat{\beta}(X - \bar{X})$,完全决定 Y 围绕均值运动的是 $X - \bar{X} = (r - \bar{r}) - (r^* - \bar{r}^*)$)。在 RE(理性预期)和风险中性假设下,上述结果有些令人费解。当然,如果我们认为那些利率比平均水平高的国家都是货币政策不稳定、政局动乱、出口软弱且产品单一(如只有少数几种农产品出口)的国家,那么在这些国家投资的高预期收益可能是系统性风险的偿付——然而,我们还必须对这种风险进行度量。我们在后面的章节中再讨论这个问题(不幸的是,我们在这方面不是很成功)。

注意,一般而言,UIP 并不意味着汇率服从鞅过程和汇率不可预测。UIP 只是意味着如果相对利率 $r - r^*$ 在所有时期中都为零,那么即期利率就是鞅($E_t s_{t+1} = s_t$)。显然,这毫无意义,因为相对利率会随着时间的演变而发生改变。也有研究表明,一组汇率变量之间的协整关系或汇率与其他宏观经济变量之间的协整关系确实意味着汇率具有可预测性(Granger - Engle 定理),但这并不必然违反市场有效性假说。正如我们已经注意到的那样,UIP 意味着通过观察今天的相对利率,我们就能预测出即期汇率的变化,但这并不意味着我们能获得超额收益。(因为在 *UIP* 条件下 $E_t R_{t+1} = 0$。)

对方程(10)进行检验的早期实证工作是通过如下方式进行的:将 $\Delta s_{t+1} - (r_t - r_t^*)$ 对 t 时刻或之前就知道的各种经济变量进行回归。Frankel(1979)发现,对 20 世纪 70 年代德国马克/美元的季度汇率数据而言,所有其他变量都不显著。Haache 和 Townend(1981)使用英镑有效汇率的月度数据(1972 年 7 月 ~ 1980 年 2 月)发现,相对利率的系数为 1,但汇率变化的滞后值和信用扩张度量指标都显著。(英镑/美元的汇率数据也能得到类似结果。)Cumby 和 Obstfeld(1981)使用 6 种主要货币对美元的周度汇率数据(1974 年 7 月 ~ 1978 年 6 月)也发现,就 6 种货币对美元的双边汇率而言,因变量 $\Delta s_{t+1} - (r_t - r_t^*)$ 滞后 16 周在统计上显著。因此,UIP 遭到拒绝。

如果我们接受 CIP,那么 UIP 的上述回归等价于将 $\Delta s_{t+1} - fp_t \equiv s_{t+1} - f_t$ 对 t 时刻已知的变量进行回归——这就是我们在下一节中要讨论的 FRU。

25.3　远期汇率无偏性

使用(9)中的抵补利率平价 $r_t - r_t^* = f_t - s_t$ 和理性预期 $s_{t+1} = E_t s_{t+1} + \varepsilon_{t+1}$，我们就得到远期汇率的无偏性(FRU)：

$$s_{t+1} = f_t + \varepsilon_{t+1} \quad \text{或} \quad E_t \Delta s_{t+1} = fp_t \tag{11}$$

$fp_t = (f - s)_t$ 在这里是远期升水。FRU + RE 通常用如下形式的方程来进行检验：

$$E_t R_{t+1} = \alpha + \delta fp_t + \gamma \Lambda_t + \varepsilon_{t+1} \tag{12}$$

或等价地，

$$E_t \Delta s_{t+1} = \alpha + \beta fp_t + \gamma \Lambda_t + \varepsilon_{t+1} \tag{13a}$$

其中，Λ_t 是 t 时刻已知的任何变量。FRU 意味着 $\alpha = \delta = \gamma = 0$ 且 $\beta = 1$，而 *RE* 则意味着 $E_t(\varepsilon_{t+1} | \Omega_t) = 0$，它包含序列无关的假设。因此，FRU + RE 意味着

$$E_t \Delta s_{t+1} = fp_t \tag{13b}$$

注意，前面的所有方程和后面的方程都可以用任意期限 k 的远期汇率来进行表示，因此，(13b)变成

$$E_t \Delta s_{t+k} = fp_t^k$$

其中，$E_t \Delta s_{t+k} \equiv E_t s_{t+k} - s_t$ 是汇率的 k 期变化，$fp_t^k = f_t^k - s_t$ 是 k 期合约中货币价格所产生的远期升水。为了简化符号，我们只考虑“单期”情形。

简单风险溢价

我们将风险溢价定义为

$$rp_t \equiv f_t - E_t s_{t+1} = fp_t - E_t \Delta s_{t+1} \tag{14}$$

现在，远期升水包含两个部分：预期贬值和风险溢价。在风险中性假设下，$f_t = E_t s_{t+1}$，因此远期市场上投机活动的期望利润为零。但是，如果有 $f_t > E_t s_{t+1}$，那么投资者在远期市场上购买外币就得支付一个溢价(高于预期价格$_{E_t s_{t+1}}$的部分)。这是引致风险的偿付，它允许交易商通过提供外币来获取预期利润。风险中性意味着 $rp_t = 0$。注意 rp_0 是风险溢价的定义，我们没有将它与任何“基础”经济变量联系起来——我们在后面会这么做。假设 CIP 成立，上式可以写为

$$\text{国外投资的收益 } R = (r_t^* + E_t \Delta s_{t+1}) - r_t + rp_t = 0$$

以 r_t 借入国内货币并投资于国外资产的套利活动在套利中止之前要求一个额外偿付 rp_t。

我们现在来考察时变风险溢价如何导致 βOLS 估计值的有偏性和不一致性。重新排列(14)并假设 *RE* 成立，检验 *FRU* 的回归方程就变成

$$E_t \Delta s_{t+1} = \alpha + \beta fp_t + \gamma \Lambda_t + (\varepsilon_{t+1} - rp_t) \tag{15}$$

如果方程(15)中的 rp_t 和相关，那么 *OLS* 估计就不是一致估计。这是计量经济学中标准的“变量误差”问题。“正确的”(即渐近无偏的)估计结果可能要用工具变量法才能获得，如果异方差存在，那么还要用 GMM 来估计协方差矩阵。

时间序列性质

从实证结果来看，即期利率是非平稳 I(1)的序列，从而 Δs_{t+1} 是 I(0)序列。RE 预测误差必须是平稳的 I(0)序列——如果不是这样，那么根据历史信息就可以对其进行预测。于是，在 FRU + RE 条件下，远期升水必须是平稳序列，s_{t+1} 和 f_t 之间的协整向量为{1，-1}。注意，FRU + RE 意味着协整向量为{1，-1}，但是，如果我们发现 s_{t+1} 和 f_t 之间本身就存在协整关系，那么这并不必然意味着 FRU + RE 成立。为了看清这一点，当 $\beta \neq 1$ 时，我们对 $\Delta s_{t+1} = \alpha + \beta(f - s)_t + \varepsilon_{t+1}$ 进行重新排列，因此，FRU + RE 的零假设不会为真。

$$(s_{t+1} - f_t) = \alpha + (\beta - 1)(f - s)_t + \varepsilon_{t+1} \tag{16}$$

如果$(f-s)$和ε_{t+1}都是平稳序列，那么$(s_{t+1}-f_t)$必然是平稳序列，因此两者之间存在协整关系。但是，当$\beta \neq 1$时，方程(16)成立，因此，协整关系并不意味着FRU成立。

使用单方程对即期汇率和远期汇率的平稳性/协整关系进行检验的文章至少有几百篇（如果没有几千篇），不同货币（相对于美元而言）的远期－即期关系都是放在VAR－ECM模型和Johansen系统中来进行检验的，协整向量一般采用$\{1,-1\}$，参见Engel(1996)的综述。这些检验的结果都不是很明确。有些研究发现，fp_t不是平稳序列，在进行单变量检验时，fp_t的非平稳性尤其明显；另外一些研究则拒绝这个序列的非平稳性，在进行面板单位根进行检验时，结论尤其明显，期限为（比方说）1个月、3个月、6个月和12个月的远期升水在Johansen VECM系统中出现（例如，Barkoulas、Baun和Chakraborty(2003)）。事实上，fp_t与单位根的“近似”程度很高。一般而言，实证结果表明$(s_{t+1}-\beta f_t)$是I(0)序列，因此s_{t+1}和f_t之间具有协整关系，但协整向量并不总是$\{1,-1\}$。那些对协整和平稳性检验都很熟悉的读者会意识到，小样本情形中的这些检验对所使用的特殊统计量尤其敏感，对“合理”备选假设的检验功效都很低（尽管检验功效比面板单位根检验好）。这些作者发现，很难相信远期升水是非平稳序列，因为我们很难相信f和s最终会离得越来越远。当然，如果我们放弃FRU＋*RE*并考虑风险溢价，如果rp_t是非平稳序列（见方程(15)，那么fp_t可能也是非平稳序列)，那么从自我反省角度来看，rp_t的非平稳性似乎同样难以令人置信（参见Lewis和Evans(1999)）。唯一合理的方法是，在对$\beta=1$进行检验时，集中精力考虑像(13)那样的方程并使用MCS或自助法来获得β的经验分布。在各种平稳性观点或fp_t的其他假设下就可以做到这一点。

检验

实证结果发现，(13a)中的β为负，对从20世纪20年代至今的各个样本期、各种货币（相对于美元而言）和远期汇率的各种期限（如1个月、3个月或1年）而言，结果都非常稳健。如果对众多研究的结果进行平均，那么β的平均值约为－1，取值范围通常为$[-4.1,-0.8]$。这就是远期升水之谜。单方程研究（如Fama(1984)、Meese和Rogoff(1983)、McCallum(1994)）的结果与此类似，单方程检验要用到多种汇率（使用SURE估计），一般要用到以$z=\{s_{t+1},fp_t\}$为变量的双边量Var模型（Baillie和McMahon(1989)、Bekaer和Hodrick(1993)）。Flood和Rose(1996)发现，β在固定汇率制度下为正（$=0.58$），但明显比1小得多，而对采用浮动汇率制的国家而言，β的值明显小于1。Bekaert和Hodrick(1992)发现，使用远期升水可以对$(s_{t+1}-f_t)$进行预测，预测的准确程度随着（远期）期限的增加而增加，就一年的期限而言，R^2可以达到30%～40%。这就否定了FRU＋RE。

注意，如果即期汇率服从随机游走，那么无论市场是否有效，β的估计值都接近于0。此外，如果即期汇率服从随机游走，那么$f_t=E_t s_{t+1}=s$，远期升水等于零，回归方程(13)中很容易产生度量误差。因此，当即期汇率接近于随机游走时，我们很难得到β准确的估计值。为了减轻这一问题的影响，我们可以将$(s_{t+1}-f_t)$对Λ_t进行回归（但是，对绝大多数货币而言，实证研究发现的结果是Λ_t的系数不等于零（使用GMM标准误差）），这再一次否定了FRU（Hansen和Hodrick(1980)）。

远期汇率f_t和s_{t+1}都是I(1)，人们一般会发现两者之间的协整关系，但这并不意味着FRU成立，也就是说，$\beta=1$（如Mcfarland、McMahon和Ngama(1994)）是无法否定FRU的，这意味着远期升水无助于预测汇率的未来变动。假设f_t和s_{t+1}之间存在协整关系，Engle－Granger定理意味着远期升水的期限结构（即$f_t^k-s_t$，期限$k=1,2,3,\cdots$）应有助于预测即期汇率的变化$\Delta_t s_{t+k}$（k期内），Clarida和Taylor(1997)在Johansen VECM框架下对这个结论进行了证明。Sarno和Valente(2004a,b)将这种方法进行了扩展，为VECM和误差项的协方差矩阵中的参数引入了三状态马尔可夫转移过程。这就是MS－VECM模型，参见第4章。就8个双边即期美元汇率(1985—2003)每周变化的单步“预测密度”而言，MS－VECM模型比线性VECM模型和随机游走模型更加接近现实。（MS－VECM模型不会进行其他预测诊断，如样本外的符号检验和其他模型所给

出的 MAE 和 RMSE。）正如在第4章所注意到的那样，我们知道汇率变动的条件分布不是正态分布，因此，如下情形并不令人感到奇怪，考虑非正态性模型的效果比不考虑这一性质的模型好。还要注意的是，这些“协整”结果（我们对其感兴趣的原因是这种动态关系表示较为简单）对有效性和 FRU 几乎没有什么直接关系。

Jensen 不等式项（JIT）

值得注意的是，如果我们考虑远期市场投机的实际收益（利润），那么 FRU + RE 就会有所差别

$$E_t\left[\frac{F_t - S_{t+1}}{P_{t+1}^{美元}}\right] = 0 \tag{17}$$

其中，$P^{美元}$是美国消费者的美元价格水平（F 用每单位外币的美元数来度量）。如果假设变量服从条件对数正态分布，那么（17）就变成

$$E_t s_{t+1} = f_t - 0.5\mathrm{Var}_t(s_{t+1}) + \mathrm{Cov}_t(s_{t+1}, P_{t+1}^{美元}) \tag{18}$$

最后两项是 Jensen 不等式项（JIT）。因此，即使对风险中性型美国投资者而言，如果我们考虑了 JIT，那么 FRU 也不成立。于是，风险中性并不意味着 UIP 成立，除非我们再次忽略 JIT。如果我们忽略协方差项（从实证结果来看，这一项很小），那么我们可以用带均值的 GARCH 模型（GARCH - in - mean）来对（18）中的系数进行估计。Bekaert 和 Hodrick（1993）及 Baillie 和 Bollerslev（1989）发现，$\mathrm{Var}_t(S_{t+1})$的系数显著不为零，从（18）中将这一项去掉也不会影响β的估计，β的估计值仍然为负。一般而言，人们认为外汇市场上的 *JIT* 很小，实际操作中可以忽略这一项。

FRU：理性预期与风险

FRU + RE 遭到拒绝的可能原因是理性预期或风险中性不成立。如果β的估计值为 -1，那么 $E_t\Delta s_{t+1} = (-1)fp_t$，$sd(E_t\Delta s_{t+1}) = sd(fp_t)$。使用 $rp_t \equiv fp_t - E_t\Delta s_{t+1}$ 和理性预期假设（预期与 t 时刻的信息（即 fp_t）独立）就可以得到

$$sd(rp_t) = 2sd(E_t\Delta s_{t+1}) \quad 且 \quad sd(f_t - E_t s_{t+1}) = 2sd(E_t\Delta s_{t+1})$$

如果 Δs_{t+1}的波动性很高，那么这意味着，为了解释风险升水之谜，我们需要一个波动性很高的风险溢价。此外，超额收益中可预测部分的方差比预期贬值本身的方差大。这两个“典型事实”对远期市场上一般均衡模型有效性的评估而言非常重要，我们在后面的章节中对其进行讨论。可以证明，

$$s_{t+1} - s_t = \alpha + \beta(f_t - s_t) + \varepsilon_{t+1}$$
$$f_t - s_{t+1} = \gamma + \delta(f_t - s_t) + v_{t+1}$$

使用（14）就可以得到

$$\beta = \frac{\mathrm{Var}(E_t\Delta s_{t+1}) + \mathrm{Cov}(rp_t, E_t\Delta s_{t+1})}{\mathrm{Var}(rp_t) + \mathrm{Var}(E_t\Delta s_{t+1}) + 2\mathrm{Cov}(rp_t, E_t\Delta s_{t+1})}$$
$$\delta = \frac{\mathrm{Var}(rp_t) + \mathrm{Cov}(rp_t, E_t\Delta s_{t+1})}{\mathrm{Var}(rp_t) + \mathrm{Var}(E_t\Delta s_{t+1}) + 2\mathrm{Cov}(rp_t, E_t\Delta s_{t+1})}$$
$$\delta - \beta = [\mathrm{Var}(rp_t) - \mathrm{Var}(E_t\Delta s_{t+1})]/\mathrm{Var}(fp_t)$$

注意，如果β的估计值为负，那么风险溢价和预期贬值率之间是一种负相关关系。根据 Fama（1984）的说法，如果 rp_t 的波动性很高，那么远期升水对即期汇率预期变化的预测结果会非常糟糕。$\delta - \beta$的值为正意味着风险溢价的方差比 $E_t\Delta s_{t+1}$的方差大，在 RE 成立的假设下，$\delta - \beta$为风险溢价中时变性的相对重要性提供了定量指导。

研究（如 Fama（1984）、Koedijk 和 Ott（1987））通常会发现 $\delta - \beta > 0$，$\delta - \beta < 0$，$\delta > 0$。Fama（1984）发现$\delta - \beta$的取值范围为1.6（对日元而言）到4.2（对比利时法郎而言）。因此，在 RE 的零假设下，FRU 命题不成立，因为（线性可加的）风险溢价具有时变性。

值得反复强调的是，上述分析的局限性在于假设潜在的时变风险溢价 rp_t 只是线性地取决于时变远期升水 fp_t，这种分析没有很强的经济理论基础。这些结果还假设 RE 成立，从而对零假设的任何违背都可以归因于时变风险溢价。我们需要一种方法来在违背 RE 和风险溢价变化之间寻找 FRU 不成立的原因。

调查数据

Frankel 和 Froot（1987）及 Froot 和 Franlkel（1989）使用投资者对未来即期汇率预期值的调查数据来说明人们应如何在 RE 不成立和风险中性不成立之间寻找零假设 FRU 不成立的原因。考虑常用的远期升水回归方程

$$\Delta s_{t+1} = \alpha + \beta fp_t + \varepsilon_{t+1} \tag{19}$$

其中，$\varepsilon_{t+1} = E_t s_{t+1} - s_{t+1} \equiv E_t \Delta s_{t+1} - \Delta s_{t+1}$是 RE 的预测误差。OLS 回归系数 β 等于

$$\beta = \mathrm{Cov}(\Delta s_{t+1}, fp_t) / \mathrm{Var}(fp_t) \tag{20}$$

将（19）中的 Δs_{t+1}代入（20）就得到

$$\beta = 1 - \beta_{RE} - \beta_{RN} \tag{21}$$

其中，

$$\beta_{RE} = -\mathrm{Cov}(\varepsilon_{t+1}, fp_t) / \mathrm{Var}(fp_t) \tag{22}$$

$$\beta_{RN} = 1 - \mathrm{Cov}(E_t \Delta s_{t+1}, fp_t) / \mathrm{Var}(fp_t) \tag{23}$$

在 RE 假设下，预测误差 ε_{t+1}与信息集合 Ω_t 独立，因此与 fp_t 独立，从而 $\beta_{RE} = 0$。此外，不管 *FRU* 中的预期是如何形成的，预期升值率 $E_t \Delta s_{t+1}$都等于远期升水 fp_t，从而 $\mathrm{Cov}(E_t, \Delta s_{t+1}, fp_t) = 1$，因此 $\beta_{RN} = 0$（即风险中性成立）。如果 RE 和风险中性都成立，那么 $\beta_{RE} = 0$ 且 $\beta_{RN} = 0$，因此，从（21）知 $\beta = 1$，这与我们所期望的相同。

如果我们拥有 $E_t s_{t+1}$的调查数据，那么我们就可以构造 $\varepsilon_{t+1} = s_{t+1} - E_t s_{t+1}$的数据序列及 β_{RE}和 β_{RN}的样本值，或等价地进行如下回归：

$$E_t s_{t+1} - s_t = \alpha + \beta_{RE}(f_t - s_t) + \omega_{t+1}$$

$$f_t - s_{t+1} = \gamma + \beta_{RN}(f_t - s_t) + v_{t+1}$$

上述方程说明，对 RE 或风险中性进行区分对 $\beta \neq 1$ 这个结果的产生的重要性。

我们要提醒一下调查数据所产生的一些问题。第一个问题是，数据是定性的（例如，被调查者的回答是“上升”、“下降”或“一样”）还是定量的（例如，被调查者回答“我预测英镑在 91 天后的汇率为 1.4 欧元/英镑”）。如果使用的是定性数据，那么进行数据变换后的各种方法对 $E_t s_{t+1}$所给出的定量结果不同。因此，对同一个预期值而言，我们可以使用不同的定量数据集合。此外，我们的定量数据要么是针对个体的数据，要么是群体平均值（或中值）。从原理上讲，*RE* 适用于个体预期，而不适用于群体平均值。

同样存在的问题是，被调查者给出的是否是经过深思熟虑后的正确答案，个体调查中的群体是固定群体还是随着时间发生变动。此外，在涉及 *FRU* 命题时，个体对 $E_t s_{t+1}$的估计时间必须与 f_t（和 s_t）的确定时间相同。最后存在的问题是，（对 $E_t s_{t+1}$的）调查数据的期限是否与 s_{t+1}已经实现的值的期限相匹配。这些问题使得我们无法根据调查数据来得出确切的研究结论。不同研究者得出的结论也不一定相同，原因可能是所使用的调查数据存在“质量差异”。

我们回到 Frankel 和 Froot（1986，1987）及 Froot 和 Franlkel（1989）的研究，他们使用的是针对美国居民的定量调查数据。他们计算出 β_{RE}和 β_{RN}后使用（21），发现 $\beta \neq 1$（实际上，β 为负），其主要原因是 RE 不成立（即 β_{RE}不等于零）。对 20 世纪 70 和 80 年代的数据而言，这一普遍性结论对 5 种以上（主要）货币和 1 个月、3 个月及 6 个月期限而言都成立。事实上，β_{RE}在统计上通常与 1 不同，尽管 Bekaert 和 Hodirick（1992）注意到回归方程中 R^2 的大大地小于 1，这似乎与风险因素的影响微不足道的情形相似。MacDonald 和 Torrance（1988）使用英国居民 1985 和 1986 年的定量调查数据获得了与 Frankel－Froot 类似的结论。

Taylor（1989b）使用的是针对美国调查者的定性调查数据，他将这些数据转换成 1981—1985

年期间的定量数据。他所发现的结果与上述结论相反，即 $\beta\neq1$ 不成立的主要原因是 $\beta_{RN}\neq0$。然而，这些证据非常微弱，因为对所考察的 4 种汇率中的 3 种而言，$\beta_{RE}=\beta_{RN}=0$，$\beta_{RN}>0$ 只在一种情形中成立。事实上，对英镑的有效汇率而言，$\beta_{RN}=1.4(t=0.15)$，但我们很难用有效汇率来解释这些结论，因为这是一"篮子"货币（每种货币都有一系列的双边远期汇率）。总而言之，使用调查数据的回归结果（参见 Froot 和 Franlkel（1989）表明，FRU 之谜的主要原因可能是系统性估计误差所致，但风险溢价也存在一定的变化。

FRU 不成立可能应归因于如下事实，代理人不是（Muth）理性的，因此他们确实会犯系统性预测错误。然而，FRU 不成立也可能归因于如下事实，代理人花时间学习新的汇率变动过程，尽管他们在进行学习还是会犯系统性错误，因为他们不知道正确的模型。如果影响汇率变动的基本因素发生持续变动，或噪声交易者对市场的影响随时间而发生变动，那么这一学习过程就会持续一段时间。换一种说法是，"比索问题"也可能存在，即使代理人是理性的，FRU 也可能不成立，因为计量经济学在一个特殊的有限样本中无法对这种预期进行正确度量。我们在下一节中讨论这些问题。

25.4　检验 FRU：Var 方法

我们可以借助于 *RE* 通过如下方法来对多期情形中的 FRU 进行检验：用已实现值 Δs_{t+m} 来替换 $E_t\Delta s_{t+m}$，然后将其对远期升水进行回归，这里的 f_t^m 是期限为 m 的远期汇率。然而，我们在这一节将使用 Var 方法来对 $E_t\Delta s_{t+m}$ 进行预测。我们已经讨论过 FRU 的单方程检验，现在我们希望将这些检验"扩展"到多时期中的远期汇率进行检验。考虑如下回归方程：

$$\Delta s_{t+1}=a_{11}\Delta s_t+a_{12}fp_t+\omega_{1,t+1} \tag{24}$$

在单期限情形下，如果远期汇率是在 $t+1$ 时刻进行交割，那么 FRU 意味着 H_0：$a_{11}=0$ 且 $a_{12}=1$。注意，EMH 也意味着预测误差与受限信息集合 $\Lambda_t=(\Delta s_t,fp_t)$ 独立。

$$\Delta s_{t+1}-E_t\Delta s_{t+1}=a_{11}\Delta s_t+(a_{12}-1)(f-s)_t+\omega_{1,t+1} \tag{25}$$

$E_t\Delta s_{t+1}=fp_t$ 在零假设下成立。在一步预测中，我们只需要方程（24）就能对 FRU 进行检验。然而，我们现在要考虑两步预测（很容易对其进行扩展）。

两期情形

假设我们拥有季度数据，但要考虑 6 个月前的远期汇率 f_t，因此 FRU 为

$$E_ts_{t+2}-s_t=E_t\Delta_2s_{t+2}=fp_t \tag{26}$$

两期预测结果为

$$E_t\Delta_2s_{t+2}\equiv E_t\Delta s_{t+2}+E_t\Delta s_{t+1} \tag{27}$$

将（24）向前滚动一期我们就可以看到，为了对 $E_t\Delta s_{t+2}$ 进行估计，我们需要对 $E_t(fp_{t+1})$ 进行预测。因此，我们需要一个方程来确定 fp，我们使用如下方程

$$fp_{t+1}=a_{21}\Delta s_t+a_{22}fp_t+\omega_{2,t+1} \tag{28}$$

方程（24）和（28）都是简单的双变量向量自回归 Var。如果 (s_t,f_t) 是 I(1) 序列，但 (s_t,f_t) 的协整参数为 $(1,-1)$，那么 $fp_t=f_t-s_t$ 是 I(0) 序列，从而 Var 中的所有变量都是平稳序列。这些平稳变量可以表示为无穷阶移动平均（向量）过程。

FRU 假说（26）意味着 Var 参数之间存在一组非线性交叉方程限制条件，这些条件确保了（26）中的两期预测误差与信息集合 $\Lambda_t=(\Delta s_t,fp_t)$ 独立。使用（27）、（24）和（28）就得到

$$\begin{aligned}E_t\Delta_2s_{t+2}&=E_t\Delta s_{t+2}+E_t\Delta s_{t+1}\\&=a_{11}(a_{11}\Delta s_t+a_{12}fp_t)+a_{12}(a_{21}\Delta s_t+a_{22}fp_t)\\&\quad+(a_{11}\Delta s_t+a_{12}fp_t)\end{aligned} \tag{29}$$

使 $E_t\Delta_2s_{t+2}$ 的表达式与 fp_t 相等就得到

$$E_t\Delta_2 s_{t+2} = \theta_1 \Delta s_t + \theta_2 fp_t = fp_t \tag{30}$$

其中，

$$\theta_1 = a_{11}^2 + a_{12}a_{21} + a_{11} \tag{31a}$$

$$\theta_2 = a_{11}a_{12} + a_{12}a_{22} + a_{12} \tag{31b}$$

显然，如果

$$\theta_1 = 0 \quad 和 \quad \theta_2 = 1 \tag{32}$$

那么(30)对Δs_t和fp_t的所有值都成立。在FRU条件下，即期汇率在t和$t+2$之间的预测误差为

$$\Delta_2 s_{t+2} - E_t\Delta_2 s_{t+2} = [\theta_1\Delta s_t + \theta_2 fp_t + \eta_{t+1}] - fp_t \tag{33}$$

其中，括号中的项由Var导出，η_{t+1}取决于$\omega_{i+1}(i=1,2)$。如果θ_1和θ_2都没有任何约束，那么就一般情形而言，预测误差的预期值取决于($\Delta s_t, fp_t$)，也就是说，取决于t时刻的信息。只有当$\theta_1 = 0$且$\theta_2 = 1$时，RE的正交性才成立。我们在上一章中曾注意到，这些约束条件可以用Wald或似然比统计量来进行检验。我们考虑似然比统计量（这个统计量在这里易于处理），将约束条件重新排列后就得到

$$a_{21} = -a_{11}(1 + a_{11})/a_{12} \tag{34a}$$

$$a_{22} = [1 - a_{12}(1 + a_{11})]/a_{12} \tag{34b}$$

将a_{21}和a_{22}代入(28)，于是带约束的Var方程为

$$\Delta s_{t+1} = a_{11}\Delta s_t + a_{12} fp_t \tag{35a}$$

$$fp_{t+1} = \left[\frac{-a_{11}(1 + a_{11})}{a_{12}}\right]\Delta s_t + \left[\frac{1 - a_{12}(1 + a_{11})}{a_{12}}\right] fp_t \tag{35b}$$

我们可以将带约束系统(35a+35b)的对数似然函数值和不带约束的系统(24)+(28)的对数似然函数值进行比较。如果对数似然函数值之间的差异很大（或很小），那么我们就要拒绝（或接受）这些约束条件。

我们很快就会说明如何用矩阵来表述上述问题（使用Var模型，滞后长度为1）及如何对其进行推广。无约束Var的矩阵形式为

$$z_{t+1} = Az_t + w_{t+1} \tag{36}$$

其中，$z_{t+1} = (\Delta s_{t+1}, fp_{t+1})$，$A(2\times 2) = \{a_{ij}\}$。令$e1 = (1,0)$，$e2 = (0,1)$，于是

$$\Delta s_t = e1'z_t \tag{37a}$$

$$fp_t = e2'z_t \tag{37b}$$

于是，

$$E_t z_{t+2} = E_t A z_{t+1} = A^2 z_t \tag{38}$$

$$E_t(\Delta_2 s_{t+1}) = E_t(\Delta_{t+2} + \Delta_{t+1}) = e1'(A^2 + A)z_t \tag{39}$$

从(26)可知，FRU意味着

$$e1'(A + A^2)z_t = e2'z_t，从而\ e2' - e1'(A + A^2) = 0 \tag{40}$$

其中，

$$A^2 = \begin{pmatrix} a_{11}^2 + a_{12}a_{21} & a_{11}a_{12} + a_{12}a_{22} \\ a_{21}a_{11} + a_{22}a_{21} & a_{21}a_{12} + a_{22}^2 \end{pmatrix} \tag{41}$$

进行“替换”之后就容易看出，(40)与我们前面在(31)中所计算出来的约束条件相同。一般来说，对期限m而言，Δs_{t+m}的远期预测结果为

$$E_t\Delta_m s_{t+m} = \sum_{i=1}^{m} E_t\Delta s_{t+i} \tag{42}$$

因此，对m期远期汇率$fp_t^{(m)}$而言，FRU为

$$E_t\Delta_m s_{t+m} = \sum_{i=1}^{m} e1'A^i z_t = fp_t^{(m)} \tag{43}$$

m期中的FRU约束条件为

$$f(A) = e2' - e1'\sum_{i=1}^{m} A^i = 0 \tag{44}$$

（43）中 Var 的预测值 $E_t\Delta_m s_{t+m}$ 给出了（m 期中）“理论上的远期升水”时间序列，正如前面的 Var 表示一样，我们可以将这个理论序列与实际远期升水（使用图表、方差比和相关系数）进行比较。

非抵补利率平价

非抵补利率平价（UIP）条件也适用于多时期情形，Var 的使用方法与上述情形完全相同。多期 UIP 条件为

$$E_t\Delta_m s_{t+m} = E_t(s_{t+m} - s_t) = (r - r^*)_t$$

例如，如果我们拥有季度数据，那么一年期债券的利差应等于汇率在连续四个季度（即 $m = 4$）中的预期变化。上述 UIP 方程与 FRU 方程类似，唯一的差别在于，右边出现的是（$r - r^*$）而不是 $fp_t^{(m)}$。显然，针对 Δs_t 和（$r_t - r_t^*$）的 Var 分析与 FRU 情形完全相同。当然，如果 CIP 成立，那么（$r_t - r_t^*$）的使用就等价于使用 fp_t 来对 FRU 进行检验。

最近的实证结果

在对 FRU（或等价地对 UIP 假说）进行检验时，近期研究所得出的一致结论是 Var 约束条件遭到拒绝（假设风险溢价具有时变特征）。对多种期限（如 3 个月、6 个月、9 个月和 12 个月）、各种信息集合、多种货币和几个样本期而言，FRU 都会遭到拒绝（例如，参见 Hakkio（1981），Baillie 和 McMahon（1989），Levy 和 Nobay（1986），Taylor（1989c））。

远期升水的期限结构

一些研究将适用于国内外利率的利率期限结构预期假说与抵补利率平价（CIP）的检验结合在一起。当然，如果你相信（根据其他证据）CIP 始终成立（我们确实会这样做！），那么这是对“EH 在两个（或多个）国家同时成立”进行检验。为了便于说明问题，考虑 3 月、6 月期的利率和远期汇率之间的抵补利率平价（CIP）关系，但我们使用的是月度数据：

$$fp_t^{(3)} = f_t^{(3)} - s_t = d_t^{(3)} \tag{45a}$$

$$fp_t^{(6)} = f_t^{(6)} - s_t = d_t^{(6)} \tag{45b}$$

其中，$d_t^{(i)} = r_t^{(i)} - r_t^{*(i)}$（$i = 3, 6$）。如果 EH 在国内和国外都成立，那么

$$d_t^{(6)} = (d_t^{(3)} + E_t d_{t+3}^{(3)})/2 \tag{46}$$

这里要用到下标 $t + 3$，因为我们使用的是月度数据。上述方程给出了远期升水的期限结构：

$$fp_t^{(6)} = (fp_t^{(3)} + E_t fp_{t+3}^{(3)})/2 \tag{47}$$

方程（47）中的概念与零息票债券及其收益率的期限结构相同，我们在第 22 章已经讨论过。显然，如果涉及 $fp_t^{(6)}$ 和 $fp_t^{(3)}$（及其他相关的信息变量）的 Var 已经给定，那么（47）就是我们熟悉的交叉方差限制条件。

已有大量的 Var 研究对（47）进行了考察，远期汇率升水的预期假说一般会遭到拒绝（例如，参见 Sarno 和 Taylor（2002））。因为我们有很强的独立证据可以证明抵补利率平价对绝大多数时期和绝大多数期限都成立，所以 Var 参数形如（47）的约束条件遭到拒绝最有可能的原因是期限结构的预期假说对国内和国外都不成立。

检验 UIP 和 EH

类似上面的分析可以用来对两个国家利率期限结构的预期假说和两国之间的 UIP 关系同时进行检验。（再次需要注意的是，如果 CIP 成立，那么对 UIP 的检验等价于对 FRU 进行检验。）对任意两个国家而言的 EH 为

$$E_t \sum_{i=1}^{n-1} (1 - i/n)\Delta r_{t+i}^j = (R_{t,n}^j - r_t^j)\ (j = 1, 2)$$

其中，r^j 是 j 国的短期利率（单期连续复利），$R_{t,n}^j$ 是 j 国的长期利率。UIP 关系为

$$E_t s_{t+1} - s_t = r_t^k - r_t^j$$

其中，s_t 等于货币 j/货币 k（的对数值）。我们现在所需要的是 Var 中的一个向量 $z_t \equiv (\Delta s_t, \Delta r_t^k, \Delta r_t^j, R_{t,n}^j - r_t^j, R_{t,n}^k - r_t^k)$，就上述三个方程而言，我们可以分别对每一个方程所隐含的交叉限制条件进行检验，也可以对三个方程同时成立时所隐含的交叉限制条件进行检验。注意，即使 Var 只取一阶滞后值，它也会涉及 25 个参数的估计（这些参数的估计值是一致的但也是有偏的）。

Bekaert 和 Hodrick(2001)使用美元与英国英镑或美元与德国马克之间的双边汇率和 1 月期、12 月期利率（月度数据，1975 年 1 月 ~ 1997 年 7 月）数据来进行上述检验。他们使用了单方程回归检验和带 Var 参数非线性约束条件的 Wald 检验，发展了拉格朗日乘子(LM)检验方法，这种检验要求在零假设下对 Var 进行带约束的估计，具体操作非常困难，但 Newey 和 McFadden(1994)说明了如何根据不带约束条件的估计值来得到带约束条件参数估计（近似）值。此外，Bekaert 和 Hodrick 对残差使用自助法得到了检验统计量的经验分布。

他们的工作对早期结论进行了证实：Wald 检验易受规模扭曲，但 LM 检验却不会。例如，在用德国 - 美国数据对三种套汇关系 Var 约束条件进行同时检验时，Wald 检验统计量和 LM 检验统计量的值分别为 47.7 和 21.3，因此，根据其中的一个检验统计量可以果断地拒绝零假设，而根据另外一个零假设却无法拒绝零假设（5% 显著性水平的临界值为 25）。但是，在 EH 和 UIP 都成立的联合零假设下，使用自助法所得到的 5% 显著性水平的临界值分别为 47.5 和 22.1，从而我们在 5% 的显著性水平可以接受这种联合约束条件。这一极其细致的研究强调经验自助法分布的使用，使用这种分布后就会发现，我们对期限结构的预期假说和 UIP 条件的拒绝比使用渐近分布的情形要更为稳妥一些。

摩擦与数据

FRU 命题假设投资者有足够的资金来使 f_t 与 $E_t s_{t+1}$ 保持一致。但现实世界中可能存在借贷限制、卖空限制和交易成本，这些都使得上述情形难以发生。有证据表明(He 和 Modest(1995)，Luttmer 和 Nishiotis(1995)，Goodhart 和 Taylor(1992))，这些因素的影响可能非常的大，但大家的一致结论是，这些因素的影响不可能大到足以"解释"数据中所可观测到的 FRU，这里的 $\hat{\beta} < 1$。

最近，Breuer 和 Wohar(1996)对 FRU 检验的数据问题进行了仔细考察，并使 t + 1 时刻的即期汇率与远期合约的交割日期完全匹配。他们也对"使用买入远期/卖出即期价格（反之亦反之）"是否对结果产生影响进行了考察，与之相比较的是使用买卖均价时的结果。例如，他们发现，在对这些结果进行精炼后，FRU 假说仍然遭到果断地拒绝。使用 1974—1993 年期间德国、英国、瑞士和日本（对美元）的一月期远期汇率数据所得到的 β 估计值多数都落在 -2.0 到 -0.6 的范围内，$\beta = 1$ 的零假设遭到拒绝的 p 值为 0.001。

25.5 比索问题与学习

在前面几节中，我们已经注意到风险中性和 RE（理性预期）的简化假设与 FRU 的实证结果不一致。在这一节中，我们将考察 FRU 在实证领域不成立的两个原因。首先，我们分析比索问题如何使得 EMH 检验结果的解释变得更加复杂，然后对学习(learning)问题进行讨论。

比索问题

FRU 在实证检验中不成立的原因可能是存在比索问题。比索问题使得研究人员无法对预期值进行正确测度，因此预测结果可能是有偏的，且与 t 时刻的信息不独立。

在 20 世纪 70 年代中期，墨西哥比索实行与美元的名义固定汇率制度。但是，因为预期未来会发生贬值（在 1976 年发生了），多年来比索的交易一直处于远期贴水状况。从表面来看，比索的远期汇率持续低于即期汇率（比如说三个月期），这一事实意味着风险中性投资者有持续的获利套利机会。

比索问题根源于如下事实，将来可能发生的事件是无法观测的（从而是无法进行量化的），但在样本数据中实际上永远也不会发生。投资者在形成预期时考虑计量经济学家所无法观察的因素是完全理性的。然而，如果在计量经济学家所考察样本中事件确实没有发生过，那么我们可能会错误地推断投资者的预期是有偏的。因此，计量经济学家可能会认为他找到了反驳 *RE* 的证据，但事实上他却没有办到。

为了进一步说明问题，我们用一种非常简单的方式来考虑比索问题。如果墨西哥政府的固定汇率政策是完全可信的（称其为“状态 1”），且政策已经维持了多年，那么对状态 t_1 下的时期 $\{t=0,1,\cdots,t_1\}$ 而言有

$$E_t s_{t+1}=s_t \tag{48}$$

因此，在预期调查数据“完全可靠”的前提下，$E_t s_{t+1}$ 是未来即期汇率 $s_{t+1}\ \{t=0,1,\cdots,t_1\}$ 的无偏估计。

现在假设外国投资者认为政府的固定汇率“承诺”已经无法兑现，比索贬值的概率为 π，维持“固定”汇率的概率为 $(1-\pi)$。我们将这一新的状态定义为“部分”可信状态。对时期 $t=\{t_1+1,t_1+2,\cdots,T\}$ 而言，理性投资者所形成的预期为

$$\begin{aligned} E_t s_{t+1} &= \pi E_t(s_{t+1}|Z2)+(1-\pi)E_t(s_{t+1}|Z1) \\ &= \pi[E_t(s_{t+1}|Z2)-E_t(s_{t+1}|Z1)]+E_t(s_{t+1}|Z1) \\ &= \pi\nabla s_{t+1}+E_t(s_{t+1}|Z1)E_t(s_{t+1}|Z1) \end{aligned} \tag{49}$$

其中，$E_t(s_{t+1}|Z1)$ 是固定汇率制（状态 1）下的汇率，$E_t(s_{t+1}|Z2)$ 是存在贬值可能性时的汇率，$\nabla s_{t+1}\equiv[E_t(s_{t+1}|Z2)-E_t(s_{t+1}|Z1)]$。因为人们在部分可信状态下预期货币会贬值，所以

$$E_t(s_{t+1}|Z2)>E_t(s_{t+1}|Z1)$$

然而，我们进一步假设墨西哥政府在部分可信状态下不会改变汇率。部分可信状态下的现实数据为 s_t（既有的固定平价）。因此，即使该时期中所收集的调查数据（其准确值为 $E_t s_{t+1}$）不等于现实值 s_t，根据（41）并使用 $E_t(s_{t+1}|Z1)=s_t$ 也可以得到

$$E_t s_{t+1}=\pi[E_t(s_{t+1}|Z2)-s_t]+s_t\neq s_t \tag{50}$$

现在假设区制转换（状态转移）确实发生了，于是，即使事后预测误差在这里不等于 RE 下的值，我们也有：

$$\begin{aligned} \varpi_{t+1} &= s_{t+1}^{(2)}-E_t s_{t+1}=[s_{t+1}^{(2)}-E_t(s_{t+1}|Z2)]-(1-\pi)\nabla s_{t+1} \\ &= \varepsilon_{t+1}+(1-\pi(1-\pi))\nabla s_{t+1} \end{aligned}$$

如果代理人确切地知道状态转移已经发生，那么 ε_{t+1} 就是 RE 预测误差。因此，事后预测误差（如果我们拥有预期的调查数据，那么它是可以观测的）不等于零，且是有偏的。此外，如果 π 随着时间变动（如经济信息发生了变化），那么 ϖ_{t+1} 对 t 时刻信息的回归一般会得到一个不等于零的系数估计值。因此，我们明显地反驳了 RE 的信息效率假设，因为预测误差与 t 时刻的信息不独立。注意，即使不可观测事件发生的概率 π 很小，如果人们认为新状态下的即期汇率存在很大的潜在变化（也就是说，∇s_{t+1} 很大），那么预测误差 ϖ_{t+1} 的“偏差”也可能会很大。

我们现在来考虑检验 FRU 时所产生的问题。在部分可信状态下，投资者认为比索很有可能会贬值，于是 $E_t(s_{t+2}|Z2)>E_t(s_{t+1}|Z1)$（记住，$s_t$ 是一美元所能兑换的比索数目，从而 s_t 增加代表的是比索的贬值）。在 FRU 条件下，远期市场的投机活动保证了

$$f_t=E_t s_{t+1} \tag{51}$$

但是，从（50）可知，如果我们仍处于状态 1 中，那么 $E_t(s_{t+1}|Z1)=s_t$，从而 $f_t>s_t$ 且 $fp_t=(f-s)_t\neq 0$。在部分可信状态下，如果 π 变动，那么 fp_t 也会变动，但如果贬值永远也不会发生，那么 $\Delta s_{t+1}=0$，即使有（51）成立，对 FRU 的回归检验也会失败。

因此，比索问题之所以会出现是因为我们是在有限的数据集合上对某个假设进行检验，一些不可观测的事件存在发生的可能但并没有发生。如果存在学习效应，那么在状态发生转移且已广为人知时 $\nabla s_{t+1}=0$，预测误差等于 RE 预测误差，于是比索问题就成为了小样本问题。当

然，如果潜在的“新”状态发生了，那么小样本问题就会再度“重现”。

根据原理，在对 $E_t s_{t+1} = f_t$ 进行检验时，我们可以通过使用预期的准确调查数据来回避比索问题。但调查数据的使用也存在一定的问题（见上）。比索问题可能是一个极为普遍的问题。显然，使用更长时期的数据可能会减轻比索问题的影响，但却无法完全将其根除。因此，对 FRU 的明显拒绝始终可以归咎于“隐含”的比索问题。

学习

如果代理人花时间去学习新的环境，那么所产生的预测误差仍然具有序列相关和非零均值特征，即使代理人使用的理性预期。假设存在两种可能的状态 $Z1$ 和 $Z2$（Levis（1989a）），于是汇率的预期值为

$$E_t s_{t+1} = \pi_t E_t(s_{t+1} \mid Z1) + (1 - \pi_t) E_t(s_{t+1} \mid Z2)$$

如果在 $m < t$ 时存在状态转移，那么代理人会使用贝叶斯规则来缓慢地更新概率值

$$\pi_t = \frac{\pi_{t-1} L(\Delta s_t, \Delta s_{t-1}, \cdots, \Delta s_{m+1} \mid Z1)}{\pi_{t-1} L(\Delta s_t, \Delta s_{t-1}, \cdots, \Delta s_{m+1} \mid Z1) + (1 - \pi_{t-1}) L(\Delta s_t, \Delta s_{t-1}, \cdots, \Delta s_{m+1} \mid Z2)}$$

其中，$L(\cdot \mid \cdot)$ 是观测数据的似然函数值，假设新旧状态都有可能发生，π_t 是不发生状态转移的后验概率。如果数据是某个状态下的数据，那么随着新状态下信息的消化和吸收，似然函数 $L(\cdot \mid Z1)$ 会减小，π_t 也趋于 0。然而，当代理人对新状态进行学习时，即使现实世界中只有一个状态会发生，他们也会为每一个状态赋予一个非零的概率。这就产生了非零的预测误差。假设现实经济变成了状态 $Z2$，那么预测误差为

$$\begin{aligned} s_{t+1}^{(2)} - E_t s_{t+1} &= [s_{t+1}^{(2)} - E_t(s_{t+1} \mid Z2)] - \pi_t [E_t(s_{t+1} \mid Z1) - E_t(s_{t+1} \mid Z2)] \\ &= \varepsilon_{t+1} - \pi \nabla s_{t+1} \end{aligned}$$

且事后预测误差不等于零。Levis（1989b）将 20 世纪 80 年代早期美元升值的一半原因归咎于代理人的贝叶斯学习效应。

考察汇率中比索问题（Levis（1988，1991））和状态转移（如 Engel 和 Hamilton（1990），Kaminsky（1993））的研究所获得的肯定结论是，20 世纪 80 年代的美元就是这种现象。问题在于，$\beta < -1$ 的 FRU 结果在多个数据时期和多种货币上出现过，令人难以置信的是存在如此之多的潜在状态转移，投资者最终不会对这些状态转移进行学习，或在这些状态转移过程中，比索问题都是局部性问题。

25.6 小结

- 外汇市场上无风险套利机会有时会在相对较长的期限（一年）中出现，但就绝大多数情形而言，大而持久的利润机会不存在，CIP 成立。
- 单方程回归、面板数据回归和 Var 方法的结果表明 FRU（和 UIP）不成立，但我们不能在风险中性不成立和 RE（理性预期）“失灵”之间寻找 FRU 遭到拒绝的原因。
- 因为远期和即期市场上可能出现中央银行频繁的干预，所以比索问题很有可能出现。然而，我们根本不可能对其进行量化，这使得解释“极为果断地拒绝 FRU 是否就意味着 EMH 遭到拒绝”变得非常困难。

第26章

外汇风险溢价建模

目的

根据 FRU 回归的实证结果来证明：$\beta < -1$（而不是 $\beta = \pm 1$）意味着，为了“拟合”事实，任何风险溢价模型中的变量都必须满足一些不等式关系。

介绍不同货币外汇收益之间关系的 CCAPM 模型 - SDF 模型、潜变量模型和仿射模型的实证检验方法。

说明如何求解和校准外汇收益的货币先行模型（cash - in - advance model），从而使由“人造经济”方法所得到的“感兴趣的统计量”与真实数据进行比较。

仅仅因为投资者持有外币资产是无法断言他们一定要获得风险溢价的，因为绝大多数这种风险都是可分散化的。只有当资产的风险与潜在风险源（如市场投资组合、消费的边际替代率）的变化之间存在协方差时，这种风险才是不可分散风险，从而要求获得额外收益。在任何情形中同时考虑国内和国外两种资产都会得出一个矛盾。如果一位美国居民认为持有欧元债券会承担汇率风险，那么欧元债券的预期收益（用美元表示）将超出美国国内安全债券的收益。但是，从欧元区居民的立场来看，他们在国内资产上获得的收益比美国（风险）资产低。因此，如果外国资产的风险溢价对美国投资者来说为正，那么对于欧元区居民来说它就为负（忽略 Jensen 不等式项）。于是，风险溢价取决于国内外资产相对风险的高低。我们使用不同模型来对时变风险溢价进行度量。

26.1　FRU 回归中 $\beta < -1$ 的含义

我们来说明，FRU 回归的实证结果 $\operatorname{plim}(\beta) < 0$ 可以用来推断变量之间的确定关系，这些关系是任何风险溢价理论都必须满足的（Fama（1984））。对本章后面与外汇风险溢价有关的 SDF 模型、仿射模型和现金先行模型的有效性评估而言，这些实证结果非常有用。FRU + RE 要求回归方程中的 $\alpha = 0, \beta = 1$ 且 $E(\varepsilon_{t+1} | \Omega_t) = 0$，

$$\Delta s_{t+1} = \alpha + \beta fp_t + \varepsilon_{t+1} \tag{1a}$$

$$\beta = \operatorname{Cov}(\Delta s_{t+1}, fp_t) / \operatorname{Var}(fp_t) \tag{1b}$$

其中，s_t 是即期汇率（国内货币/外币）的对数值，$fp_t \equiv f_t - s_t$，f_t 等于远期汇率的对数值。我们将风险溢价定义为：

$$rp_t \equiv f_t - E_t s_{t+1} = fp_t - E_t \Delta s_{t+1} \tag{2}$$

在 RE 下，

$$\Delta s_{t+1} = E_t \Delta s_{t+1} + \varepsilon_{t+1} \tag{3}$$

ε_{t+1}与 t 时刻的信息独立，于是

$$\mathrm{Cov}\{fp_t,\Delta s_{t+1}\}=\mathrm{Cov}\{fp_t,E_t\Delta s_{t+1}\} \tag{4a}$$

使用(4a)并将(2)中的 $E_t\Delta s_{t+1}$代入就得到

$$\mathrm{Cov}\{fp_t,\Delta s_{t+1}\}=\mathrm{Var}(fp_t)-\mathrm{Cov}\{fp_t,rp_t\}$$

将协方差项中的fp_t用(2)式中的结果代入就得到

$$\mathrm{Cov}\{fp_t,\Delta s_{t+1}\}=\mathrm{Var}(fp_t)-\mathrm{Cov}\{E_t\Delta s_{t+1},rp_t\}-\mathrm{Var}\{rp_t\} \tag{4b}$$

将(4b)代入(1b)就得到

$$\mathrm{plim}(\hat{\beta})=1-\beta_{rp} \tag{5}$$

其中，

$$\beta_{rp}=\frac{\mathrm{Cov}\{E_t\Delta s_{t+1},rp_t\}+\mathrm{Var}(rp_t)}{\mathrm{Var}(fp_t)} \tag{6}$$

实证研究发现 $\mathrm{plim}(\hat{\beta})<0$，从而$\beta_{rp}>1$，由(6)就可以得到

$$\mathrm{Cov}\{E_t\Delta s_{t+1},rp_t\}+\mathrm{Var}(rp_t)>\mathrm{Var}(fp_t) \tag{7}$$

但根据(2)，有，

$$\begin{aligned}\mathrm{Var}(fp_t)&=\mathrm{Var}(E_t\Delta s_{t+1}+rp_t)\\&=\mathrm{Var}(E_t\Delta s_{t+1})+\mathrm{Var}(rp_t)+2\mathrm{Cov}\{E_t\Delta s_{t+1},rp_t\}\end{aligned} \tag{8}$$

由(7)和(8)可知，$\beta_{rp}>1$ 意味着

$$\mathrm{Cov}\{E_t\Delta s_{t+1},rp_t\}+\mathrm{Var}(E_t\Delta s_{t+1})<0 \tag{9}$$

因为 $\mathrm{Var}(E_t\Delta s_{t+1})>0$，所以由(9)可以得到我们的第一个不等式，即即期汇率的预期变化与风险溢价的变化(即投机性远期利润$f_t-E_ts_{t+1}$)负相关：

$$\mathrm{Cov}\{E_t\Delta s_{t+1},rp_t\}<0 \tag{10}$$

其中，$rp_t\equiv f_t-E_ts_{t+1}$。由(9)可知，我们要求

$$|\mathrm{Var}(E_t\Delta s_{t+1})|<|\mathrm{Cov}\{E_t\Delta s_{t+1},rp_t\}|=|\rho\sigma(E_t\Delta s_{t+1})\sigma(rp_t)| \tag{11}$$

因为$|\rho_{\max}|=1$，所以上式给出了我们的第二个不等式，即风险溢价的方差大于即期汇率预期变化的方差：

$$\mathrm{Var}(rp_t)\equiv\mathrm{Var}(f_t-E_ts_{t+1})>\mathrm{Var}(E_t\Delta s_{t+1}) \tag{12}$$

我们在后面还会提到这两个不等式。

26.2 CCAPM 模型

CCAPM 模型可以用来对外汇市场上的投机性收益进行建模。注意，在即期市场上进行国外投资的名义超额收益为 $R_{s,t+1}\equiv(\Delta s_{t+1}+r^*-r)$，其中 r^* 是国外利率，r 是国内利率。在外汇市场上进行投机活动且收益为 $R_{F,t+1}=s_{t+1}-f_t$ 是可能的。但是，如果 CIP 成立，那么 $f_t=s_t+r_t-r_t^*$ 且 $R_{s,t+1}=R_{F,t+1}$。在 CAPM 进行检验时；我们可以使用两种不同的方法来对投机收益进行度量。对任意两种资产 i 和 j 而言，一阶条件(FOC)或欧拉方程为：

$$U'(C_t)=\theta E_t\{R_{i,t+1}U'(C_{t+1})\}=\theta E_t\{R_{j,t+1}U'(C_{t+1})\} \tag{13a}$$

重新排列(13a)并使用欧元和美元收益就得到

$$0=E_t\{(R_{t+1}^{美元}-R_{t+1}^{欧元})M_{t+1}\} \tag{13b}$$

其中，

$R^{美元}$ = 美元资产的实际收益

$R^{欧元}$ = 欧元资产的实际收益

$M_{t+1}=\theta U'(C_{t+1})/U'(C_t)=\theta(C_{t+1}/C_t)^{-\gamma}$(幂效用函数情形)

S = 汇率(=国内货币/国外外币，美元/欧元)

C_t = 国内(美国)的实际消费

CCAPM 模型使用的是实际收益，所以对任意货币 j 而言，方程(13b)可以写成：

$$0 = E_t\left\{\left(\frac{(F_{j,t} - S_{j,t+1})P_t^{\text{美元}}}{S_{j,t}R_{t+1}^{\text{美元}}}\right)M_{t+1}\right\} \tag{14}$$

其中，$P_t^{\text{美元}}$ 是美元价格水平，(14)背后的直觉是，远期合约要求在 t 时刻没有投资，从而 $(F_t - S_{t+1})/S_t$ 是远期投机收益。注意，$U'(C_{t+1})/P_{t+1}^{\text{美元}}$ 是每一美元的边际效用。在(13a)中，代表性投资者的偏好与他所处的位置无关。在联合对数正态分布和幂效应函数的联合假设下，对任意货币而言，由(14)可以得到(见 Kaminsky 和 Periga(1990)，p. 54)：

$$\begin{aligned} E_t\Delta s_{j,t+1} - fp_{j,t} = & -0.5\text{Var}_t(\Delta s_{j,t+1}) - \text{Cov}_t(\Delta s_{j,t+1}, \pi_{t+1}^{\text{美元}}) \\ & + \gamma\text{Cov}_t(\Delta s_{j,t+1}, \Delta c_{t+1}) \end{aligned} \tag{15}$$

其中，$\pi_{t+1} = \ln(P_{t+1}/P_t)$。注意，只有汇率项才有下标 j，因为价格和消费都是对国内经济(投资者)而言的。因为 $rp_t = fp_t - E_t\Delta s_{t+1}$，所以“风险溢价”只取决于前面两个 Jensen 不等式项，而“真正的”风险溢价实际上就是 $+\gamma\text{Cov}_t(\Delta s_{j,t+1}, \Delta c_{t+1})$(当 $\gamma = 0$ 时这一项等于 0)。

对 CCAPM 模型的检验是基于(14)那样的一阶条件或(15)那样的显式解(在对数正态分布假设下)。后者的困难在于需要对时变条件协方差进行度量。如果协方差项的数目大于 2，那么这项工作就尤为困难(见下)。

Mark(1985)、Hodrick(1989)和 Modjtahedi(1991)在幂效用函数假设下针对多种货币(相对于美元)的 1 月、3 月和 6 月远期汇率使用美国(加总)的人均消费(非耐用品和服务)数据来对(14)进行了估计。与 γ 有关的(过度识别)交叉方程约束条件遭到了拒绝，$\hat{\gamma}$ 的值为 40 ~ 70，这位于我们可接受的范围之外。以英镑为计价货币时使用英国的消费数据会得到类似的结论。

Kaminsky 和 Periga(1990)使用德国马克、日元和英镑(对美元)的汇率数据对方程(15)进行了估计，他们假设“RE 残差”服从多变量 GARCH 模型(参见 Engle 和 Kroner(1995))。所使用的 5 个变量是：

$$z'_{t+1} = [\Delta c_{t+1}^{\text{美元}}, \pi_{t+1}^{\text{美元}}, (s_{t+1} - f_t)^j, (s_{t+1} - f_t)^k, (s_{t+1} - f_t)^m] \tag{16}$$

其中，j、k 和 m 分别为日元、英镑或马克(对美元的汇率)。五个方程为

$$z_{t+1} = A_0 + A_1(L)z_t + Dvec(H_{t+1}) + v_{t+1} \tag{17}$$

H_{t+1} 是误差项的协方差矩阵，$vec(H_{t+1})$ 是用误差项协方差矩阵的元素所堆积成的向量：

$$\begin{aligned} vec(H_{t+1}) = \{ & h(1,1), h(1,2), \cdots, h(1,5), h(2,2), h(2,3), \cdots, h(2,5)\}, \\ & h(3,3), \cdots, h(3,5), h(4,4), h(4,5), h(5,5) \end{aligned}$$

方程(15)对 15 × 15 矩阵 D 所施加的约束条件是：

$$\begin{aligned} & d(3,3) = d(4,4) = d(5,5) = \gamma; \\ & d(3,10) = d(4,13) = d(5,15) = -1; \\ & d(3,7) = d(4,8) = d(5,9) = -0.5, \\ & d(i,j) = 0, \text{当 } i \text{ 和 } j \text{ 取其他值时。} \end{aligned}$$

非零的 $\{d_{ij}\}$ 值“识别”出了(15)中条件方差和协方差项对三种货币均值 $(\Delta s_{t+1} - fp_t)$ 的效应。他们没有拒绝 $\text{Var}_t(\Delta s_{j,t+1})$ 的三个系数为 −0.5 和 $\text{Cov}_t(\Delta s_{j,t+1}, \pi_{t+1}^{\text{美元}})$ 的三个系数为 −1 的联合假设。对三种货币而言，相对风险规避系数为常数。然而，根据 t 时刻的信息是无法对残差 v_{t+1} 进行预测的。但他们却发现，对日元和英镑(而不是马克)而言，可以使用远期升水来对残差进行预测，因此，模型无法对所有的“风险溢价”$f_t - E_t s_{t+1}$ 进行解释。绝大多数人指责的是 γ 的估计值大于 350。因此，远期市场上投机的事前收益 $f_t - E_t s_{t+1}$ 的波动性是如此的高以至于无法用 $\text{Cov}(\Delta s_{t+1}, \Delta c_{t+1})$ 来对其进行解释，除非 γ 的值很大(对股权溢价而言，CCAPM 模型所给出的结果与之类似)。Backus、Gergory 和 Telmer(1993)在效用函数中使用的是习惯消费 $c_t - \lambda c_{t-1}$，这一项在 M_{t+1} 中会出现，但仍然无法改变模型的解释困境。

外汇收益的 SDF 模型(CCAPM 模型)

Smith 和 Wickens(2002)沿着 Kaminsky 和 Periga(1990)的路线使用 CCAPM 模型对外国投资

收益进行估计，但他们使用的是更为简单的 GARCH 过程，相关系数为常数。从本质来看，这允许他们以单方程的形式来估计 GARCH 过程，而不是在多变量模型中使用极大似然法来估计 GARCH 过程（在参数化程度非常高的情形中估计非常困难）。协方差为 $\rho_{ij}\sigma_{i,t}\sigma_{j,t}$，时变标准差可以由单变量 GARCH 过程来产生。收益方程要使用这些协方差，我们再次可以用但方程技巧来对其进行估计。

对美国国内投资者而言，国外投资超额收益的 CCAPM 模型为：

$$E_tR_{t+1}+0.5V_t(R_{t+1})=k_{美国}\mathrm{Cov}_t(\Delta c_{t+1}^{美国},R_{t+1})+\mathrm{Cov}_t(\Delta p_{t+1}^{美国},R_{t+1}) \tag{18}$$

其中，$R_{t+1}=(\Delta S_{t+1}+r_t^*-r_t)$。即期汇率 s_t 用国内货币/外国货币来度量（这里为美元/英镑）。因此，美元贬值的速度与美国消费增长率和通货膨胀率之间的协方差越大，对美国投资者而言，国外债券的风险溢价就越高。对英国国内投资者而言，类似方程为

$$E_tR_{t+1}-0.5V_t(R_{t+1})=-k_{英国}\mathrm{Cov}_t(\Delta c_{t+1}^{英国},R_{t+1})-\mathrm{Cov}_t(\Delta p_{t+1}^{英国},\Delta c_{t+1}^{英国}) \tag{19}$$

将上述两个方程相加就可以消去 $V_t(\cdot)$ 项，ER_{t+1}取决于英美国两国消费和通货膨胀率之间的协方差：

$$\begin{aligned}E_tR_{t+1}=0.5\{&k_{美国}\mathrm{Cov}_t(\Delta c_{t+1}^{美国},R_{t+1})+\mathrm{Cov}_t(\Delta p_{t+1}^{美国},R_{t+1})\\&-k_{英国}\mathrm{Cov}_t(\Delta c_{t+1}^{英国},R_{t+1})-\mathrm{Cov}_t(\Delta p_{t+1}^{英国},\Delta c_{t+1}^{英国})\}\end{aligned} \tag{20}$$

使用 1975.1 ~ 1997.12 期间美元 - 英镑汇率的月度数据发现，CCAPM 模型中的协方差项在统计上都不显著，相对风险规避系数的符号是错的且值很大。后来，Smith 和 Wickens（2002）通过假设货币模型中的变量会对 SDF 产生（线性）影响来对模型进行了扩展。收益与如下变量之间的协方差现在出现在预期收益方程中：消费、产出和货币增长。方程变为：

$$R_{t+1}=\gamma_1R_t+\gamma_2fp_t+\gamma_3V_t(R_{t+1})+\phi^{美国}Z_{t+1}^{美国}+\phi^{英国}Z_{t+1}^{英国}+\varepsilon_{t+1} \tag{21}$$

为了得到一般的模型，这里将 R_t、fp_t 和 $V_t(R_{t+1})$加了进来，但是，如果 SDF 模型成立，且 $Z_{t+1}^{美国}$和 $Z_{t+1}^{英国}$分别表示美国和英国的协方差项，那么我们预期 $\gamma_i=0(i=1,2,3)$。人们在平均（预期）收益方程中发现了一些支持条件协方差的证据，但这些证据的说服力不是很强。此外，与模型相对比的是，滞后收益和远期升水在统计上仍然显著。因此，SDF 模型中的“货币因子”只能获得很小的支持，远期升水之谜仍然存在（Mark 和 Wu（1998）及 Engel（1996）得出了相似的结论）。

人们经常在多变量 GARCH 框架中对条件方差进行建模。但是，即使我们在考虑了几个协方差项，待估参数的个数也可能很大（参见第 29 章），除非我们对 GARCH 模型的参数进行限制。但这似乎会限制 GARCH 方法在 SDF 资产定价模型中的作用。

潜变量和跨期 CCAPM 模型

潜变量模型通过假设条件协方差项是一组可观测变量 z_t 的线性函数来将 CCAPM 模型中的条件协方差项进行替换。对任意资产 j 而言，CCAPM 模型可以表示为

$$E_t(R_{j,t+1}-R_{0,t+1})=\frac{-\mathrm{Cov}_t(M_{t+1},R_{j,t+1})}{E(M_{t+1})} \tag{22}$$

其中，R_0 为资产收益，它与 M_{t+1}之间的协方差为零（R_0 扮演着与国内 CCAPM 模型中无风险利率相同的角色）。现在将 $R_{b,t+1}$当作均值方差前沿收益，它是 $R_{0,t+1}$和最小方差收益的加权平均。于是，我们可以证明（Hodrick（1987））：

$$E_t(R_{j,t+1}-R_{0,t+1})=\beta_{j,t}E_t(R_{b,t+1}-R_{0,t+1}) \tag{23}$$

其中，$\beta_{j,t}=\mathrm{Cov}_t(R_{b,t+1},R_{j,t+1})/\mathrm{Var}_tR_{b,t+1}$。

尽管 $\beta_{j,t}$可能具有时变性，但 $\beta_{j,t}/\beta_{i,t}$等于常数（$=\lambda$）也是合理的，从而预期相对收益为常数，$E_t(R_{i,t+1}-R_{0,t+1})/E_t(R_{j,t+1}-R_{0,t+1})=\lambda$。（从宽泛意义上将，这与“国内”CAPM 模型等价，资产 i 和 j 之间的相对收益取决于 β 值的相对大小。）

我们现在来说明模型所隐含的限制条件，如果假设资产 i 和 j 之间的相对收益线性地取决于一组可观测变量 z_{kt}，那么

$$R_{i,t+1}-R_{0,t+1}=\sum_{k=1}^{n}\alpha_k z_{kt}+u_{i,t+1} \tag{24a}$$

$$R_{j,t+1}-R_{0,t+1}=\sum_{k=1}^{n}\delta_k z_{kt}+u_{j,t+1} \tag{24b}$$

潜变量模型意味着对 $k=1,2,\cdots,n$ 而言，约束条件 $\alpha_k/\delta_k=\lambda$ 成立。检验通常将（$S_{j,t+m}-F_{j,t+m}$）/$S_{j,t}$当作远期市场上货币 j 的投机性收益（对任意期限而言，如 $m=1$ 或 3 个月），并将它当作因变量。变量 z 可能包括远期贴水和历史估计误差。好几个研究（如 Hansen 和 Hodrick（1983）、Campbell 和 Clarida（1987）、Huang（1989）、Lewis（1990））都发现了与交叉方程约束条件不一致的证据，但期限为三个月的模型效果比期限为一个月的模型效果更好。Cumby 和 Huizinga（1992）对模型进行了更为深入的检验，认为使用不同货币对（18）进行回归所得到的拟合值高度相关（如果估计误差不存在，那么它们应完全相关）。他们发现这些相关系数的取值范围为 0.33 ~ 0.65，且与 1 在统计上存在显著差异，因此我们应拒绝模型。

Giovanni 和 Jorion（1987）、Mark（1988）及 Cumby（1988）试图使用方程（17）中 β 具有时间变化的各种统计模型，总体而言，他们无法拒绝约束条件，这为这些模型提供了支持证据。然而 Engel（1996，*p.* 163）认为，拒绝零假设的失败是因为检验功效太低，也就是说，“模型”和“替换模型”都不令人满意（即“解释”效果很差），因此施加约束条件不会使“拟合”效果降低很多。

26.3　外汇收益的仿射模型

与期限结构的仿射模型一样，我们可以用仿射（线性）结构来为国内外经济提供强制性变量 z_{1t}和 z_{2t}。我们在一阶条件中使用这些变量来生成利率变化（即即期汇率和远期升水的变化）方程。为了获得总体想法，我们考虑一个单因子模型，模型中的因子相互独立、因子具有国家特殊性且与随机折现因子（SDF）m_{t+1}线性相关。在 CIR（Cox、Ingersoll 和 Ross）结构中，

$$-m_{t+1}=z_{1t}+\lambda\sigma\sqrt{z_{1t}}\varepsilon_{1,t+1} \tag{25}$$

$$-m^{*}_{t+1}=z_{2t}+\lambda\sigma\sqrt{z_{2t}}\varepsilon_{2,t+1} \tag{26}$$

$$z_{i,t+1}-\mu=\theta(z_{it}-\mu)+\sigma\sqrt{z_{it}}\varepsilon_{i,t+1}\quad(i=1,2) \tag{27}$$

为了简单起见，我们假设参数 λ、σ 和 θ 对所有国家而言都相同。在这种情形中可以证明，对每一个国家而言，短期利率与 z 线性相关：

$$r_{i,t}=(1-0.5(\lambda\sigma)^2)z_{it}\quad(i=1,2) \tag{28}$$

外汇变量由下面的式子给出：

$$fp_t=[1-0.5(\lambda\sigma)^2](z_{1t}-z_{2t}) \tag{29a}$$

$$E_t\Delta s_{t+1}=(z_{1t}-z_{2t}) \tag{29b}$$

$$f_t-E_t s_{t+1}=-0.5(\lambda\sigma)^2(z_{1t}-z_{2t}) \tag{29c}$$

如果在 FRU 方程中有 $\beta<1$，那么我们知道这要求风险溢价和即期汇率的预期贬值之间存在负相关关系。但是，由（29b）和（29c）可以很容易地看出，在我们的仿射模型中风险溢价 $rp_t=f_t-E_t s_{t+1}$与预期贬值率负相关。然而，我们的模型认为，在 $E_t\Delta s_{t+1}$对 fp_t 的 FRU 回归中 β 的值为 $1/[1-0.5(\lambda\sigma)^2]$，这个值为正且大于 1，参见（29a）和（29*b*）。但这与 $\beta<1$ 的经验证据矛盾，因此，我们的 CIR 单因子仿射模型遭到了拒绝。Backus、Foresi 和 Telmer（2001）使用的是两因子模型，因子之间相互依赖，他们还对模型进行了一些参数化处理，这使得 β 的值为负，但代价是出现了不可观测因子和远期升水的一些不可接受的分布特征。简而言之，SDF 仿射模型到目前为止还没有对远期升水之谜进行很好的解释。

26.4 FRU 与货币先行模型

在货币先行模型（CIA）中，CCAPM 模型的标准一阶条件（FOC）和购买力平价（PPP）关系都成立，但多了一些不等式约束条件，即代理人有足够多的国内货币或国外货币，从而在任何国家中进行消费都可以达到合意的水平。这是一种禀赋经济。在偏好对消费具有时间可分离性标准货币先行模型中，远期汇率行为的如下典型事实试图得到解释：

（ⅰ）远期贴水具有持续性，但波动性现对较低。

（ⅱ）预期的远期投机利润$f_t - E_t s_{t+1}$比远期升水的波动性更高。

（ⅲ）即期汇率比$(f_t - E_t s_{t+1})$或fp_t的波动性更高，且近似为鞅过程。

（ⅳ）FRU 被违背，因为在$\Delta s_{t+1} = \alpha + \beta fp_t + \varepsilon_{t+1}$中我们发现$\hat{\beta} \approx -1$，而它本应该为 +1。

注意，预期的远期投机利润$f_t - E_t s_{t+1}$通常用$f_t - s_{t+1}$对远期升水进行回归的拟合值来进行度量。记住，我们在前面就发现系数β可以写成

$$\beta = \frac{\mathrm{Cov}_t\{E_t \Delta s_{t+1}, f_t - E_t s_{t+1}\} + \mathrm{Var}(E_t \Delta s_{t+1})}{\mathrm{Var}_t(fp_t)} \tag{30}$$

可以证明，负的β值要求：

（a）$\mathrm{Cov}_t\{\cdot\} < 0$ （见方程 10）

（b）$\mathrm{Var}_t(f_t - E_t s_{t+1}) > \mathrm{Var}_t(E_t \Delta s_{t+1})$ （见方程 12）

标准货币先行模型

在标准 CIA 模型中，投资者追求跨期可分离效用的最大化，效用函数是两个国家家庭消费的函数。为了能在每一个国家购买消费品，投资者必须持有每个国家的货币（现金）。一阶条件为：

$$S_t = \frac{U'_{2t}/P_t^2}{U'_{1t}/P_t^1}$$

$$q_t^j = \theta E_t\left[\frac{U'_{j,t+1}/P_{t+1}^j}{U'_{j,t}/P_j}\right] \text{（为国家 1 或 2）}$$

$$\frac{F_t}{S_t} = \frac{q_t^2}{q_t^1}$$

其中，U'是边际效用，P^j是国家$j = (1,2)$中的商品价格，q_t^j是债券价格，S_t为即期汇率，F_t为远期汇率，θ为折现因子。第一个方程是 PPP 条件，实际汇率等于国内外商品的边际替代率。第二个方程给出了国内和国外债券的名义价格q_t^j，第三个方程是 CIP 关系。因为货币数量理论成立，所以价格与国内货币供给成比例，从而即期汇率的变化取决于两个国家的相对货币供给增长。由债券定价方程（和实际冲击与名义冲击之间的协方差为零的假设）可知，远期贴水$\ln(F_t/S_t)$也取决于相对的货币增长。然而，相对货币增长具有持续性，所以标准 CIA 模型认为即期汇率和远期贴水的变化都具有持续性。但在数据中，即期汇率的变化近似为白噪声，从而标准 CIA 模型无法解释这一典型事实。

Moore 和 Roche（2002）通过假设所有资产组合决策都是在货币冲击之前进行的来对标准 CIA 模型进行扩展，这使得债券价格之比能对实际汇率（和边际替代率）产生影响：

$$S_t = E_t\left(\frac{U'_{2,t+1}/P_{t+1}^2}{U'_{1,t+1}/P_{t+1}^1}\right)\left(\frac{q_t^1}{q_t^2}\right)$$

从而模型认为，一方面即期汇率有着较高的波动性，另一方面即期汇率具有一定的持续性，因为债券价格的波动性变化能使（持续的）相对货币供给的影响所有减弱。假设货币和消费增长服从 AR（1）过程，与“标准”模型相比，校准后的模型可以减小即期汇率的自相关性，增加它的波动

性，但校准模型的拟合结果仍然与数据有着很大的差别。

标准货币先行模型能解释上面所说的(*i*)或(*ii*)，但不能对两者都进行解释，它也不能解释 $\beta<0$ 时 FRU 为什么不成立(Bekaert(1996))。

习惯持续

Moore 和 Roche(2002)在标准 CIA 模型中加入了 Campbell 和 Cochrane(2000)的习惯持续思想，以此来考察这些"典型事实"是否可以被负值，修改后的模型能对(ⅰ)和(ⅱ)进行解释，但无法对(ⅲ)进行解释。更为仔细地考察这种方法是值得的，因为它为"校准"和"人造经济"模型提供了一个"案例研究"。我们不会"逐行"导出这个模型，因为多数结论是 Campbell 和 Cochrane(1999)类模型的两国版本，我们已对此进行了讨论。

总的来说，这种方法的包括如下步骤：

(a) 建立带约束的最优化问题，导出系统内生变量(如消费增长、资产收益、货币供给)之间关系的一阶条件。

(b) 选取一些变量作为外生变量，用一个简单的随机过程来对其进行建模，模型基本上能拟合数据的时间序列特征。例如，在禀赋经济中，假设消费增长是均值为常数的 *niid*。这就是"校准"。

(c) 在此人造经济中对外生变量进行模拟，使用一阶条件来导出内生变量(如即期汇率、远期汇率、利率)的时间路径，计算"感兴趣的统计量"。一般而言，统计量是诸如重要变量的均值、方差和相关系数的"矩"，但可以将其扩展到偏斜度、峰度和回归关系(如即期汇率对变化对远期升水的回归)上去。

(d) 重复"(c)"$m(=10000)$次，计算感兴趣统计量的平均值，将其与真实数据进行对比。如果"人造经济"中的统计量与真实数据中的结果非常接近，那么可以认为模型取得了成功。

我们也可以改变强制变量的时间序列性质(如用 AR(1)来对消费增长进行建模)，重复整个过程来决定感兴趣的统计量对这些"合理"选择的敏感性。

模型中有两个国家、两种商品和两种货币，代表性投资者的习惯偏好型跨期幂效应函数为：

$$\sum_{t=0}^{\infty}\theta^{t}U(C_{it}^{1},C_{it}^{2})=\sum_{t=0}^{\infty}\frac{\theta^{t}(X_{it}^{1})^{1-\gamma}}{1-\gamma}+\frac{\theta^{t}(X_{it}^{2})^{1-\gamma}}{1-\gamma}\tag{31}$$

其中，$X_{it}^{j}=(C_{it}^{j}-H_{it}^{j})/C_{it}^{j}$，下标 j 表示国家 1 或 2，i 表示家庭 1 或 2。C_{it}^{j}是 i 国家庭消费 j 国的商品数量，H_{it}^{j}是习惯消费，X_{it}^{j}是剩余消费，$X_{it}^{j}=0$ 表示可能的最坏状态。货币先行模型的约束条件为，j 国货币 N^{j} 要(被本国居民或外国居民)用来购买 j 国的商品。

$$N_{it}^{j}\geqslant P_{t}^{j}C_{it}^{j}\quad i=1,2,j=1,2\tag{32}$$

P_{t}^{j} 是用 j 国货币表示的 j 国价格。对正的利率而言，(32)是一个等式。在 t 期期末($t+1$ 期的开始时刻)，国内家庭对国内货币的持有量为

$$N_{1,t+1}^{1}\geqslant P_{t}^{1}C_{t}^{1}+B_{1t}^{1}-F_{t}G_{t}^{1}\tag{33}$$

第一项是从消费禀赋中获得的收益，第二项是因为名义债券的赎回而产生的，第三项是长期持有的外币合约数量($G_{t}^{1}>0$)，F_{t} 是单期远期汇率(国内货币/外国货币)。国内的外币持有量为

$$N_{1,t+1}^{2}\geqslant B_{1t}^{2}+G_{t}^{1}\tag{34}$$

外国家庭对国内和国外货币持有量的方程与之类似：

$$N_{2,t+1}^{1}\geqslant B_{2t}^{2}+F_{t}G_{t}^{2}\quad(\text{国内本币})\tag{35}$$

$$N_{2,t+1}^{2}\geqslant P_{t}^{2}C_{t}^{2}+B_{2t}^{2}-G_{t}^{2}\quad(\text{外国货币})\tag{36}$$

其中，$G_{t}^{2}>0$ 是外国在远期汇率上的空方头寸。商品、货币和远期市场的均衡条件为

$$C_{t}^{j}=C_{1t}^{j}+C_{1t}^{j}\quad j=1,2\tag{37a}$$

$$N_{t}^{j}=N_{1t}^{j}+N_{2t}^{j}\quad j=1,2\tag{37b}$$

$$G_{t}^{1}=G_{t}^{2}\tag{37c}$$

国际风险完全共担，从而均衡消费是当前禀赋的一半，$C_{jt}^{i}=0.5Y_t^i$，Y_t^i 是 i 国家在 t 时刻的禀赋数量。通常的一阶条件为

$$E_t\left[\theta(1+r_t^j)\frac{U'(C_{t+1}^j)P_t^j}{U'(C_t^j)P_{t+1}^j}-1\right]=0\quad(j=1,2)\tag{38}$$

其中，对 $j=(1,2)$ 而言，我们有 $U'(C^j)=(C^jX^j)^{-\gamma}$，$r_t^j(j=1,2)$ 是两国的名义利率。另外一个一阶条件就是 PPP：

$$\frac{S_tP_t^2}{P_t^1}=\frac{U'(C_t^2)}{U'(C_t^1)}=\frac{(C_t^2X_t^2)^{-\gamma}}{(C_t^1X_t^1)^{-\gamma}}\tag{39}$$

其中，S_t 为即期汇率，远期汇率由下式给出：

$$\frac{F_t}{S_t}=\frac{(1+r_t^1)}{(1+r_t^2)}\tag{40}$$

模型中的强制变量是消费增长和货币增长，消费增长服从 *niid*，

$$\frac{C_{t+1}^j}{C_t^j}=(1+\mu_{t+1}^j)\tag{41}$$

$$\mu_{t+1}^j=\bar{\mu}^j+v_{t+1}^j\qquad v_{t+1}^j\sim N(0,\sigma_{vj}^2)\qquad(j=1,2)\tag{42}$$

货币增长的均值服从 AR(1)过程

$$\frac{N_{t+1}^j}{N_t^j}=(1+\pi_{t+1}^j)\tag{43}$$

$$\pi_{t+1}^j=(1-\rho)\bar{\pi}^j+\rho\pi_t^j+u_{t+1}^j\qquad u_{t+1}^j=N(0,\sigma_{uj}^2)\tag{44}$$

假设实际冲击 v_t 和货币性冲击 u_t 之间的相关系数为 0（在基准模型中），不同的国家参数取值都相同。与 Campbell 和 Cochrane（1999）一样，剩余消费服从“持续的”AR(1)过程，调整参数为 ϕ，$\lambda(x)$是剩余消费比率的非线性函数

$$x_{t+1}^j=(1-\phi)\bar{x}^j+\phi x_t^j+\lambda(x_t^j)v_{t+1}^j\qquad(j=1,2)\tag{45}$$

与标准货币先行模型相比，这一关系对“新”的模型预测至关重要。非线性函数 $\lambda(x_t^j)$ 可以放大任何冲击 v_{t+1}对实际消费增长的影响，$\phi>0.9$ 传递了这种持续性。

可以证明（基本上是根据(38)），实际利率 rr（在两个国家中）为常数，（对每一个国家而言）名义利率为

$$r_t^j=rr^j+[(1-\rho)\bar{\pi}+\rho\pi_t^j-\mu^j]-[0.5\sigma_n^2+(1-2\gamma)[1-2\gamma(x_t^j)]\sigma_v^2)]\tag{46}$$

在(46)中使用$f_t-s_t=r_t^1-r_t^2$ 就可以得到远期升水：

$$\begin{aligned}fp_t&=[\rho(\pi_t^1-\pi_t^2)]+[\gamma\sigma_v^2\{\lambda(x_t^1)-\lambda(x_t^2)\}]\\&\approx[\rho(\pi_t^1-\pi_t^2)]-[\gamma(1-\phi)(\bar{X}/\gamma)(x_t^1-x_t^2)]\end{aligned}\tag{47}$$

(47)中的最后一项是对 $\lambda(x_t)$进行线性化后的结果。将(32)中的货币先行等式代入 PPP 方程(39)，取对数后再将消费、货币和剩余消费的外生过程代入后就得到：

$$\Delta s_{t+1}=[\rho(\pi_t^1-\pi_t^2)+(u_{t+1}^1-u_{t+1}^2)+(1-\gamma)v_{t+1}^2]+[\gamma(\Delta x_{t+1}^1-\Delta x_{t+2}^2)]\tag{48}$$

预期远期利润为“(47)减去(48)的期望值”，可以证明它等于

$$f_t-Es_{t+1}=\gamma(1-\phi)\left(1-\frac{\bar{X}}{\gamma}\right)(x_t^1-x_t^2)\tag{49}$$

在方程(46)、(47)和(48)中，如果我们设 $\lambda(x_t^j)=0$，那么模型就变为标准货币先行模型（无习惯偏好），因此，从本质来看，$\lambda(x_t)$给出的是一种另外的“风险溢价”。

在(46)中，中括号中的第一项等于货币增长减实际消费增长，可以将其视为预期通货膨胀。当剩余消费较低时，$\lambda(x_t)$较高且名义利率必须高到足以补偿通货膨胀所导致的风险增加。在(47)中，$\rho(\pi_t^1-\pi_t^2)$项是通货膨胀率（即货币增长）之差，习惯$(x_t^1-x_t^2)$对fp 的影响主要取决于国内剩余消费比率的最初水平（相对于国外剩余消费比率而言）。在方程(48)中 Δs_{t+1}新增的“部分”取决于剩余消费增长的差异，这就是为什么模型中的即期汇率存在高度波动的原因。注意，在标准模型中，$E_t\Delta s_{t+1}$仅由通货膨胀率差$(\pi_t^1-\pi_t^2)$决定。

我们现在来就算模型中关键变量的方差和协方差，假设剩余消费在稳定状态下的值为

$$\bar{X}^j = \sigma_{vj}\sqrt{\gamma/(1-\phi)} \tag{50}$$

在(40)、(47)、(48)和(49)所得到的远期贴水、即期汇率和预期远期利润的无条件方差为：

$$\mathrm{Var}(fp_t) = 2\left(\frac{\rho^2\sigma_u^2}{1-\rho^2} + \gamma(1-\phi)\sigma_v^2\sigma_x^2\right) \tag{51a}$$

$$\mathrm{Var}(\Delta s_{t+1}) = 2\left(\frac{\sigma_u^2}{1-\rho^2} + (1-\phi)^2\gamma^2\sigma_x^2 + \sigma_v^2\left(1-\frac{\gamma}{X}\right)^2\right) \tag{51b}$$

$$\mathrm{Var}(f_t - E_t s_{t+1}) = 2(1+\phi)^2\gamma^2\left(1-\frac{\bar{X}}{\gamma}\right)^2\sigma_x^2 \tag{51c}$$

σ_x^2 是 $x_t^j(j=1,2)$ 的无条件差。习惯模型在标准模型中增加了 σ_x 和 $\gamma/\bar{X}$ 两项，这可以看作是增加了即期外汇收益、远期贴水和预期远期利润的绝对波动率大小(当 $\bar{X}<1$ 时，$\gamma/\bar{X}>2$)。相对波动率使得即期收益的波动性(51b)要么比远期贴水的波动率(51a)，要么比预期远期利润的波动率(51c)增加得更快一些。但预期远期利润的波动率比远期贴水的波动率要增加得更快一些。这些定性的波动率关系式与实际数据中所发现的结果相吻合，但我们还不清楚这些结果是否与实际数据中的真实矩相匹配。可以证明，“习惯”增加了远期贴水的持续性，但也减小了即期收益的持续性(对 $\phi>\rho$ 且 σ_v^2 相对较小的情形而言)。

我们现在回到 FRU 回归(1a)中的 β 系数上去。

$$\beta = \frac{\mathrm{Cov}(E_t s_{t+1} - s_t, f_t - E_t s_{t+1}) + \mathrm{Var}(E_t\Delta s_{t+1})}{\mathrm{Var}(fp_t)}$$

对 β 为负的情形而言，我们要求协方差为负且绝对值比 $\mathrm{Var}(E_t\Delta s_{t+1})$ 大。第二个条件(Fama 1984)等价于 $\mathrm{Var}(f_t - E_t s_{t+1}) > \mathrm{Var}(E_t\Delta s_{t+1})$。

在货币先行模型中，$E_t\Delta s_{t+1}$ 和 $f_t - E_t s_{t+1}$ 都取决于 $x_t^1 - x_t^2$，即国内和国外剩余消费比率对数值之差。但 Δs_{t+1} 比 $f_t - E_t s_{t+1}$ 对习惯更敏感，后者的变动主要受 $E_t s_{t+1}$ 所主导，这就产生了如下负相关关系，$\mathrm{Cov}(E_t s_{t+1} - s_t, f_t - E_t s_{t+1}) < 0$，这是 $\beta<0$ 的必要条件。不幸的是，模型也预言 $\mathrm{Var}(E_t\Delta s_{t+1}) > \mathrm{Var}(f_t - E_t s_{t+1})$，这意味着在习惯模型中有 $\beta>1$，但这与经验数据完全不一致。模型增加了即期收益 Δs_{t+1} 的波动性，主要原因是预期即期收益的波动性增加了。为了使 $\beta<0$，我们 $\mathrm{Var}(\Delta s_{t+1}) > \mathrm{Var}(f_t - E_t s_{t+1}) > \mathrm{Var}(E_t\Delta s_{t+1})$，在习惯模型中最后一个不等式不成立。

以上都是一些解析结果，但我们可以用多变量模型来对外生的随机过程进行模拟，计算相应的矩，将结果与真实数据进行对比。Moore 和 Roche(2002)发现，对英国、美国和日本的双边汇率而言，蒙特卡罗“矩”和经验矩之间存在非常好的对应关系。例如，对美国 - 英国而言，人造经济模型给出的即期收益的波动性、预期远期利润和远期贴水分别为 9.5%、1.09% 和 1.04%。另外，对“人造经济”而言，即期收益的持续性(即 AR(1) 系数)为 -0.01，远期升水的持续性为 0.66，而 FRU 回归中 β 值为 1.79(Moore 和 Roche(2002)，表 6)。对其他双边汇率而言，使用较准模型所得到的(FRU 回归)β 的估计值落在 0.96 ~ 2.1 的范围内，但这与真实数据中的结果(约为 1 或比 1 小)相去甚远。因此，为了“拟合” 数据，模型需要一种机制来产生较高的无条件反差 $\mathrm{Var}(\Delta s_{t+1})$ 和 $E_t\Delta s_{t+1}$ 的适度波动性(换句话说，即期收益的波动率高得令人惊讶)。不幸的是，即使我们考虑时变风险溢价，模型也无法解释 FRU 之谜。

实际和名义汇率

尽管 CIA 模型无法“解释”远期升水之谜，但基本上能拟合如下典型事实：(i)名义汇率和实际汇率的波动性都很高，两者之间高度相关；(ii)实际汇率具有一定的持续性，名义汇率近似服从随机游走。

标准 CIA 模型(即效用取决于两个国家的消费水平)和粘性价格假设可以解释上述典型事实(Chari、Kehoe 和 McGrattan(2001)，Bergin 和 Feenstra(2001))。粘性价格假设考虑了实际汇率的持续性和名义汇率的近似随机游走行为(受相对货币增长所驱使)。这些模型需要一个很大

的 γ 值（即跨期替代弹性很小）来产生名义汇率的高波动性（参见 $x_t^i = 1$ 时的方程(31)，这是效用函数的一个标准设定）。

即使假设价格完全富有弹性，如果我们像方程(45)那样认为习惯形成的持久性导致了实际汇率的持续性，那么 *CIA* 模型也无法解释上述典型事实。为了看清这一点，注意由如下式子给出的实际 S^r 和名义 S：

$$S_t^r = \frac{U'(C_t^2)}{U'(C_t^1)} = \frac{(C_t^2 X_t^2)^{-\gamma}}{(C_t^1 X_t^1)^{-\gamma}}$$

$$S_t = \frac{S_t^r P_t^1}{P_t^2}$$

取对数后再使用 $P_t^1 = P_t^2 = (n^1 - n^2) - (y^1 - y^2)$，就得到，

$$S_t^r = (y_t^1 - y_t^2) + \gamma(x_t^1 - x_t^2)$$

$$S_t = (\gamma - 1)(y_t^1 - y_t^2) + \gamma(x_t^1 - x_t^2) + (n_t^1 - n_t^2)$$

这就是完美国际风险共担，从而均衡消费是当前禀赋的一半，$C_{jt}^i = 0.5Y_t^i$，这里的 Y_t^i 是 i 国在 t 时刻的禀赋总量。与标准效用函数相比，习惯持续模型在实际汇率和名义汇率的决定因素中引入了剩余消费比率 x_t（的对数值）(Moore 和 Roche(2002))。如果没有习惯持续项 x_t，那么我们需要比 1 更大的 γ 值来产生实际汇率和名义汇率之间的正相关和名义即期汇率的强波动性(Chari、Kehoe 和 McGrattan(2001))。通过引入习惯持续项 x_t（它具有很强的持续性($\phi \approx 0.97$)），我们不需要价格粘性或大的 γ 值就能获得这些典型事实(Moore 和 Roche(2002))。名义汇率的近似随机游走行为部分是因为相对消费（即禀赋）和相对货币供给增长（如果使用的是基础货币序列，如广义货币，这样的货币序列有着很强的自回归性，那么货币供给增长的持续性会很弱）具有近似随机游走行为。

使用粘性价格和关于消费的标准效用函数的 CIA 方法，或使用弹性价格和习惯持续效用函数的 CIA 方法，在拟合实际汇率和名义汇率时间序列性质（如前二阶矩）的典型事实方面取得了一定的成功。然而要注意的是，这些模型将消费（和货币）的时间序列过程视为给定来推断汇率的性质，这些模型都是禀赋模型而不是一般均衡模型。在计价“货币”的选择方面存在一定的弹性，这为研究人员“拟合事实”给予了一定的一些灵活性，因为可能的规范检验方法（如对参数进行限制）有很多。但是，正如上面提到的那样，CIA 模型和其他方法一样无法解释远期升水之谜。

26.5 小结

- 在使用用一阶条件或直接对条件方差进行建模或使用潜变量方法来进行检验时，CCAPM 模型在多种货币和多个时期上都表现得很差。
- CIA 模型有着与 CCAPM 模型一样的一阶条件，但需要用国内和国外货币来购买消费品进行消费。外汇收益等数据可以用“人造经济”来产生。为了对模型进行“检验”，我们要使用“人造数据”来生成感兴趣的“统计量”（如方差和协方差），然后将其与真实数据的“统计量”结果进行比较。
- CIA 模型能解释远期汇率和即期汇率的一些典型事实，但无法解释 $\beta < 0$ 时的 FRU 结果。事实上，习惯持续 CIA 模型能产生 $\beta > 1$ 的“人造经济”。

第27章

汇率与基本经济因素

目的

- 说明相对货币供给如何通过PPP关系来影响即期汇率。这是弹性价格货币模型(flex - price monetary model,FPMM)的基础。
- 分析存在粘性价格和精明投机者的模型如何产生汇率超调。这就是粘性价格货币模型(sticky - price monetary model,SPMM)。
- 说明非抵补利率平价(UIP)关系如何导致前瞻性货币模型(forward - looking monetary model)。
- 使用协整技巧、动态误差修正模型(ECM)和样本外预测绩效等来考察货币模型的实证检验。
- 勾画“新开放经济宏观经济学”(new open - economy macroeconomics)的主要特征。

27.1　货币模型

人们使用基于“基本经济因素”的各种模型来分析即期汇率的变动,我们在这一章中只对主要思想进行概述。“各种形式的货币模型主导了汇率的理论和实证研究文献”可能是正确的,我们将讨论几个这样的模型,如弹性价格货币模型、粘性价格货币模型和Frankel实际利率模型。正如我们即将看到的那样,这些模型对汇率变动的解释还远非成功。事实上,经济学家还没有对影响汇率的基本经济因素达成共识,这也是政策制定者为什么要通过诸如布雷顿森林协定和欧洲的ERM来限制汇率波动的部分原因,后一种情形的目标是走向完整的货币联盟,即欧洲货币联盟(EMU)。弹性价格货币模型(FPMM)侧重于经常账户而不是资本账户,假设价格富有弹性且产出由经济的供给方面外生决定。在浮动汇率体制下,FPMM认为货币的快速增长和汇率贬值之间有着非常密切的关系(反之亦然)。比如说,这基本上与意大利、英国、德国和日本在20世纪70年代前几年的情形和一些拉丁美洲国家在20世纪七八十年代的情形相一致。事实上,根据模型的预测结果,教科书中的Mundell - Fleming模型在充分就业假设下会得出与FPMM相似的结论。

不幸的是,FPMM不能很好地解释小型开放经济中实际汇率(即价格竞争力)的大幅波动,如20世纪70年代前几年和80年代早期在英国、新西兰和意大利发生的情形。FPMM把“货币”当成仅有的重要资产,所以忽略了国际收支平衡表中资本账户的其他资产流动。一旦我们意识到资本流动的重要性,在外汇管制逐步取消后这种重要性明显增加,我们就必须强调预期问题。投机性短期资本流动是对国内外相对利率的反应,但也取决于对汇率变动的预期。粘性价格模

型（SPMM）借助理性预期假说来处理汇率预期，它通常假设非抵补利率平价（UIP）成立。商品市场上的价格调整很缓慢，且由超额需求通过附加价格预期的菲利普斯曲线（PEAPC）来决定。粘性价格和高资本流动性意味着货币（和财政）政策的变化可能导致名义汇率和实际汇率“大”波动，也可能导致汇率超调。

汇率文献中反复出现的一个主题是讨论汇率对国内利率变化的反应。FPMM 认为国内利率的上升会导致货币贬值，而 SPMM 却得出相反的结论。实际利率货币模型（real interest rate monetary model，RIMM）对汇率与利率之间的关系进行了阐述。

最后，SPMM 的缺陷在于它的隐含假设：国内外资产具有完全替代弹性，这使得模型无法分析由经常账户不平衡所导致的存流量交互作用，但这可以用汇率的组合平衡模型（PBM）来进行弥补。

弹性价格货币模型

FPMM 依赖于 PPP 条件和稳定的货币需求函数。假设（对数）货币需求取决于（对数）实际收入 y、价格水平 p 和（债券）利率水平 r，假设“国外”货币需求函数与之类似。国内和国外的货币均衡为：

$$m^s = p + \phi y - \lambda r \qquad (1a)$$

$$m^{s*} = p^* + \phi^* y^* - \lambda^* r^* \qquad (1b)$$

国外变量在这里用星号来进行标记。在 FPMM 中，国内利率是外生的，这是一个很奇特的性质。这个假设意味着国内利率与外生的世界利率之间存在刚性联系，因为我们假设“资本完全流动”且汇率变化的预期值为零。如果假设产出是充分就业水平下的产出（即新古典供给曲线），那么任何超额货币都只能一对一地影响“富有完全弹性”的国内价格水平，因此“货币中立性”成立。

当共同货币的价格达到了均衡时，简而言之，当 PPP 成立时：

$$s = p - p^* \qquad (2)$$

贸易品“市场”（即经常账户）也会达到均衡。世界价格 p^* 对国内经济而言是外生的，它由世界货币供给来决定，s 是（对数）即期汇率（国内货币/国外货币）。国内货币供给决定了国内价格水平，从而汇率由相对货币供给决定。将（1）代入（2）并进行重新排列和代数计算后就得到：

$$s = (m^s - m^{s*}) - \phi y + \phi^* y^* + \lambda r - \lambda^* r^* \qquad (3)$$

（3）中蕴含的传递机制是：（i）国内货币供给的增加导致对国外商品（和资产）需求的增加，从而对外币产生超额需求并使国内汇率贬值；生产者于是在国内价格上进行“套利”，促使价格上升到新的贸易商品价格水平时为止。（ii）超额货币导致国内商品超额需求的产生，随后经由菲利普斯曲线，国内价格出现上涨。随之而来的便是需求转向为相对便宜的国外商品，导致国内汇率产生向下的压力。

值得注意的是，在 FPMM 中，产出或国内利率的变化对汇率的效应与凯恩斯模型所发现的结论相反。在 FPMM 中，更高的产出水平或更低的国内利率都会导致国内货币需求的增加。在后一种情形中，为了使货币市场达到均衡，我们需要一个更低的利率水平，这会导致汇率的升值（例如，参见 Frenkel、Gylfason 和 Helliwell（1980），Gylfason 和 Helliwell（1983））。现在，可能会出现名义利率的上升，因为货币政策出现了紧缩，或预期通货膨胀率 π 出现了上升（Fisher 假说）：

$$r = \psi + \pi \qquad (4)$$

这里的 ψ 是实际利率。将此关系加入到 FPMM（3）中后我们就会看到，高预期国内通货膨胀率与高名义利率和国内汇率的快速贬值（即 s 取“更大”的值）相伴而生。因此，当我们将 Fisher 假说和 FPMM 合在一起时，我们称其为超通胀－FPMM，利率与汇率之间的关系就有些不合常理。出现超通胀－FPMM 这一术语的原因是 r 由超通胀 π 的变化所主导（如 20 世纪 20 年代在德国发生的情形）。这种称呼非常恰当，但我们可能更倾向于认为，s 的变化值而不是它的水平值取决于预期通货膨胀率，这与我们在后面要讨论的 Frankel（1979）“实际利率”模型一样。

FPMM 的检验可以通过估计汇率形如（3）的方程或考察 PPP 关系和货币需求函数的稳定性

来进行。就方程(3)而言，对 20 世纪 70 年代早期即浮动汇率时期的众多双边汇率来说，根据它所得到的实证结果非常合理(见 Bilson(1978))，但对 20 世纪 70 年代后期而言，除那些有着高通货膨胀的国家(如阿根廷和巴西)外，它表现得很糟糕。20 世纪 70 年代资本流动性的增加可以解释 FPMM 不成立的原因，汇率在这里只取决于相对价格(PPP)。尽管对 PPP 关系进行检验存在种种困难，我们还是已经指出它只在长期中才成立。另外，已有大量文献对 20 世纪 70 年代到 90 年代 G10 经济中货币需求函数的不稳定性进行了描述(见 Cuthbertson(19991a,b,1997))，这意味着 FPMM 不成立。

在 20 世纪 70 年代的后五年中，FPMM 无法对小型开放经济的汇率行为进行准确描述。例如，对 1979—1981 年期间的英国而言，英镑的名义有效汇率(即相对于一篮子货币而言的汇率)大幅升值，尽管英国当时的货币供给比“世界”货币供给的增长还要快一些。然而，更令人吃惊的是，实际汇率(即价格竞争力)在这一时期中升值了约 40%，但在 1981—1984 期间却贬值了相同的幅度。FPMM 只能通过国内(和国外)价格水平的变化对相对货币供给的短期不同滞后反应来使实际汇率的变动。面对上述各种证据，这些滞后变量可变性很强，换句话说，FPMM 不能很好地解释这种现象。

SPMM：Dornbusch 汇率超调

SPMM 对汇率超调(Dornbusch(1976))和实际产出的短期变动给出了解释。模型能解决在 FPMM 中所发现的问题：一个违反直觉的结果是国内利率的上升导致国内货币贬值。在 SPMM 模型中，如果名义利率出现了未曾预料的上升，从而实际利率也会上升，那么汇率贬值这一传统结论就可能成立。

与 FPMM 一样，SPMM 在某种意义上说是一种“货币主义”观点，因为借助产出垂直的新古典供给曲线(或等价地垂直的长期菲利普斯曲线)就可以看出长期货币中立性成立。然而，PPP 只在长期中才成立，因此实际净贸易平衡的短期变化是允许的。SPMM 的关键之处在于如下假设：传统的稳定货币需求函数和 UIP 成立。假设外汇市场上的投资者能对汇率的未来走势形成理性预期：他们立即对新信息产生反应，这使得汇率发生“跳跃”，出现频繁变动。此外，在 SPMM 中，资本账户和货币市场在所有时期中都“出清”，但商品市场因为价格粘性而不能“出清”。“弹性价格”市场与“固定价格”市场的结合就可以产生汇率超调。

我们简单地叙述 Dornbusch(1976)模型，首先描述主要的行为性假设，然后分析紧缩性货币政策对经济的影响。

UIP 关系表示的是资本账户的均衡条件。外汇投机商投资国外市场的期望百分比收益为 $r^* + \mu$，其中，r^* = 国外利率，μ = 外币的预期升值率(本币贬值率)。在资本完全流动和风险中性的假设下，均衡条件要求

$$r = r^* + \mu \qquad (5)$$

假设在开始的时候有 $r = r^* = 10\%$，因此 $\mu = 0$。假设汇率预期是退化预期或回归预期(regressive expectation)。如果真实汇率位于长期均衡汇率 $\bar{s}$ 之下，那么投资者会预期真实汇率朝着长期汇率方向上升，也就是说，国内货币的即期汇率趋于贬值：

$$\mu = \theta(\bar{s} - s) \quad (0 < \theta < 1) \qquad (6)$$

如果 Dornbusch 模型中的其他方程已经给定，那么这个“预期方程”与理性预期完全一致，因为这个递归公式允许事后对预期进行修正。货币市场均衡意味着

$$m^s = m^d = -\lambda r + \phi y + p \qquad (7)$$

商品市场上的总需求(AD)为

$$AD = \delta(s - p + p^*) - \sigma r + \gamma y + \gamma' \qquad (8)$$

第一项表示实际汇率对净贸易量(即出口减进口)的影响，第二项是富有弹性的实际(投资)支出($-\sigma r$)，第三项是将支出对进口的效应扣除后的消费函数(γy)，最后一项是外生的需求因子(γ')，如政府支出等。“供给方面”用垂直的长期菲利普斯曲线来表示。通货膨胀率对商品市场

上的超额需求有所反应；价格朝着均衡价格进行缓慢调整（$0<\Pi<1$）：

$$\Delta p=\Pi(AD-\bar{y})=\Pi[\delta(s-p+p^{*})-\sigma r+\gamma y+\gamma'-\bar{y}] \tag{9}$$

这里的 $\bar{y}$ 是充分就业水平下的产出。

弹性价格：长期观点

我们从商品和货币市场均衡开始讨论，假设 *UIP* 成立（即 $r=r^{*}$，$\mu=0$）。考虑货币供给减少 1% 的情形。如果价格具有完全弹性，那么价格水平下降 1% 就能使货币市场重新恢复均衡（长期利率水平没有变化）。另外，如果汇率升值 1%，那么实际汇率仍然为常数，实际总需求继续与总供给一致。在长期中，利率不会变动，因此实际投资也不会变动。在利率不变的条件下，我们与其未来汇率仍然为常数，且 UIP 成立。在货币紧缩 1% 后，价格会下降 1%，汇率将升值 1% 来维持价格竞争力（PPP）。这就是“长期”结果。

固定价格：短期超调

与上述情形相对照的是，我们现在假设价格和产出在短期内具有粘性。由于 y 和 p 具有“粘性”，货币供给要求 r 上升来使货币市场“出清”（$dr=-(1/\lambda)dm^{s}$），如方程（7）所示。r 上升导致潜在的资本流入，只有当人们预期国内汇率贬值从而 UIP 重新区域成立时资本流入才会停止。但是，根据方程（6），国内货币的预期贬值要求实际即期汇率立即升值到它高于长期均衡值之上的水平。因此，汇率相对于长期值而言处于“超调”状态。

今天的实际升值背后的经济学原理如下。如果下降 1%，且 $\lambda=1/2$，那么利率将上升 2% 来使货币市场出清（y 和 p 在短期内是固定的）。如果人们对即期汇率的预期没有变化，那么较高的利率将导致资本流入国内经济（此时 UIP 不再成立）。国内货币需求的增加将推高今天的实际即期汇率。在国外投机商停止购买国内货币之前即期汇率要上升多少呢？有利于国内经济的利率差异为 2%。因此，今天的即期汇率会要上升到比已知的长期值高 2% 的水平，因为理性投机者会认为来年的汇率会下降 2%。国内货币 2% 的预期下降恰好抵消国内利率 2% 的增加，现在 UIP 成立。因此，如果即期汇率最初为 100，那么在国内利率上升 2% 后，即期汇率值在缓慢下降到年末的 101 长期值（即长期升值率为 1%）之前会立即攀升到 103。因此，汇率相对于长期值而言处于超调状态，因为它直接从 100 直接跳到 103，而不会处于长期值 101 的水平，它不会从 100 单调地上升到 101。所有这些都非常漂亮且一致，因为投资者都是理性的，知道正确的经济模型（至多有一组白噪声）。

简单地陈述这些结果背后的数学知识是很有用的。由于菲利普斯曲线是垂直的，产出处于长期的固定水平，货币中立性意味着 $d\bar{p}=dm^{s}$。由于 PPP 在长期中也成立，所以 $\mathrm{d}\bar{s}=\mathrm{d}\bar{p}=\mathrm{d}m^{s}$（变量上的横线表示长期值）。转向短期情形，假设 p 和 y 固定，从而货币市场的任何短期非均衡都只有通过 r 来进行调整：

$$\mathrm{d}r=-\mathrm{d}m^{s}/\lambda \tag{10}$$

为了保持短期 UIP 的关系，汇率的预期升值 μ 必须等于利率的微分 $\mathrm{d}r$（注意 $\mathrm{d}r^{*}=0$）：

$$\mathrm{d}\mu=\mathrm{d}r=-\mathrm{d}m^{s}/\lambda \tag{11}$$

根据预期方程（6），使用（11）就可以得到汇率的短期变化

$$\mathrm{d}s=\mathrm{d}\bar{s}-\mathrm{d}\mu/\theta=[1+(\theta\lambda)^{-1}]\mathrm{d}m^{s} \tag{12}$$

因为 $\theta\lambda>0$，所以即期汇率的初始变化 $[1+(\theta\lambda)^{-1}]\mathrm{d}m^{s}$ 会高于“单位”长期变动：$\mathrm{d}\bar{s}=\mathrm{d}m^{s}$。“超调”的部分原因显然是货币政策强制运行的渠道受到了限制。在开始的时候，货币市场上的所有调整都是通过利率来进行的，只有在长期价格才会达到使货币市场达到均衡的水平，利率也才会回到初始水平。尽管从上述分析无法明显地看出这一点，但风险中性假设对汇率超调而言是至关重要的。注意，与 FPMM 预言相对照的是，汇率对利率的反应与人们的直观预期相同：利率预料不到的跳跃（这是货币供给减少的结果）导致了国内货币的升值。

Frankel 实际利率货币模型

Frankel(1979)给出了一个用于分析利率变动对汇率影响的一般化模型，我们将其称作“实际利率货币模型”(Real Interest Monetary Model, RIMM)。这个模型给出了针对名义利率的 Dornbusch 关系($\partial s/\partial\pi > 0$)和针对预期通货膨胀率的超通胀 FPMM 关系($\partial s/\partial\pi > 0$)。汇率也可能超调长期均衡值。

Frankel 模型假设存在非抵补套利，但对 Dornbusch 的汇率预期方程进行了修改，他增加了($\pi - \pi^*$)项来反映长期的相对预期通货膨。“预期方程”为

$$s^e - s = \theta(\bar{s} - s) + (\pi - \pi^*) \tag{13}$$

由 UIP 关系就可以得到

$$s^e - s = r - r^* \tag{14}$$

预期贬值率($s^e - s$)取决于汇率对其均衡值的偏离，正如我们所知道的那样，这就给出了 Dornbusch 类型的结果。另外，如果 $s = \bar{s}$，那么预期贬值率由国内和国外货币之间的预期通货膨胀差异来决定：正如我们即将看到的那样，这一项产生了超通胀的 FPMM 结果。Frankel 认为预期方程与理性预期是一致的。(我们将不涉及这一方面。)结合(13)和(14)并进行重新排列后我们就得到

$$s - \bar{s} = (1/\theta)[(r - \pi) - (r^* - \pi^*)] \tag{15}$$

即期汇率围绕均衡值的运动由相对利差决定。在长期均衡中，$s = \bar{s}$，这意味着 $\bar{r} - \bar{r}^* = \pi - \bar{\pi}^*$；中括号一项可以写成$[(r - r^*) - (\bar{r} - \bar{r}^*)]$。只有在紧缩性货币政策使名义利差($r - r^*$)上升到高于相对通货膨胀所给出的它的长期水平($\bar{r} - \bar{r}^*$)时，“当前”汇率才会升值到高于长期均衡的水平($\bar{s} - s > 0$)。

我们现在假设 PPP 在长期中成立，假设货币需求函数是通常的货币需求函数(为了简化起见，取 $\phi = \phi^*, \lambda = \lambda^*$)，我们就得到长期的表达式(与 FPMM 中的结果一样)：

$$\begin{aligned} \bar{s} &= \bar{p} - \bar{p}^* = \bar{m} - \bar{m}^* - \phi(\bar{y} - \bar{y}^*) + \lambda(\bar{r} - \bar{r}^*) \\ &= (\bar{m} - \bar{m}^*) - \phi(\bar{y} - \bar{y}^*) + \lambda(\pi - \pi^*) \end{aligned} \tag{16}$$

我们在这里使用了 $\bar{r} - \bar{r}^* = \pi - \pi^*$(“国际 Fisher 效应”，它隐含在超通胀 FPMM 中)。Frankel 模型的关键部分是预期方程(13)和汇率长短期决定因素之间的差异。将(15)中的 $\bar{s}$ 代入(16)就得到 Frankel(“简约式”)汇率方程：

$$\begin{aligned} s &= \bar{m} - \bar{m}^* - \phi(\bar{y} - \bar{y}^*) - (1/\theta)(r - r^*) + [(1/\theta) + \lambda](\pi - \pi^*) \\ &= \bar{m} - \bar{m}^* - \phi(\bar{y} - \bar{y}^*) + \alpha(r - r^*) + \beta(\pi - \pi^*) \end{aligned} \tag{17}$$

其中，$\alpha = -(1/\theta)$，$\beta = (1/\theta) + \lambda$。现在，我们可以根据参数 α 和 β 来刻画三个相互竞争的模型。

从表 1 就可以明显看出，在 Frankel 模型中，如果在通货膨胀预期保持不变时提高利率，那么我们就得到了 Dornbusch 类型的结果($\partial s/\partial r < 0$)。这种情形很有可能与货币供给出现预料不到变化的情形相符合，货币供给预料不到的变化会对利率产生直接影响(为了使货币市场“出清”)，但不能认为这种影响是长期的，因此它不会对 π 产生影响。另一方面，名义利率 r 和通货膨胀预期 π 的等幅增长会导致汇率贬值($\beta + \alpha > 0$)，这是一个 FPMM 类型的结果。因此，通过在 Dornbusch 模型(即(13))中加入辅助假设，预料之中的货币供给增加就可能导致国内货币发生预期的真实贬值。Frankel 模型隐含地强调汇率对预料到的和没有预料到的货币供给和利率的反应差异。

表 1　　*Frankel* 实际利率模型

模型	参数
Frankel	$\alpha<0,\beta>0;\ \lvert\beta\rvert>\lvert\alpha\rvert$
FPMM	$\alpha>0,\beta=0$
FPMM - 超通货膨胀	$\alpha=0,\beta>0$
Dornbusch - SPMM	$\alpha<0,\beta=0$

组合平衡模型

FPMM 和 SPMM 有两个重要的简化假设:所有国内和国外资产都具有完全替代性;经常账户盈余或赤字的任何财富效应都可以忽略。汇率组合平衡模型考察将这些假设进行放松的结果(例如,参见 Branson(1977),Isard(1978),Dornbusch 和 Fischer(1980))。组合平衡模型(PBM)中的汇率水平,至少在短期内是由所有金融资产(如货币、国内外债券)的市场供给和需求来决定的。在 PBM 中,经常账户盈余(赤字)表示国外资产的国内净持有量出现了增加(减少)。后者会影响国内财富水平,从而影响合意的资产需求,这会最终对汇率产生影响。因此,PBM 是包含经常账户、资本积累速度和价格水平之间交互作用的内生汇率调整动态模型。在最简单的 PBM 中,国内货币 M、国内债券 B 和国内居民持有的国外债券 B^* 都具有传统的需求函数:

$$M/W=m(r,r^*+E_t\Delta s_{t+1})$$

$$B/W=B(r,r^*+E_t\Delta s_{t+1})$$

$$SB^*/W=B^*(r,r^*+E_t\Delta s_{t+1})$$

其中,r = 国内利率,r^* = 国外利率,$E_t\Delta s_{t+1}$ = 国内货币的预期贬值。国内货币、债券和国外债券可以相互替代。国内财富由国内货币、债券和国债券所组成,国外债券的变动由经常账户盈余 $T(S/P)$ 所控制,因为资本账户在均衡状态下必须与经常账户相等:

$$W=M+B+SB^*$$

$$\dot{B}^*=T(S/P)+r^*B^*$$

贬值改善了贸易平衡 $T'>0$。模型可以通过静态预期假设或理性预期假设来进行求解,汇率超调是可能的。注意,商品价格在模型中无法确定(通常假设长期中性(long - run neutrality)成立)。"理论"PBM 忽略了对权益资产的需求,但这种需求在分析公开市场操作对汇率的影响时就显得非常有用。

由于数据的限制,对 PBM 进行检验是很困难的。直接检验涉及对资产需求函数进行估计(如 Lewis(1988),Cuthbertson 和 Galindo(1999))。另外一种方法是对 PBM 进行求解,这能得出汇率的简化式方程,方程涉及货币之外的资产,即国内债券和净国外债券(通常用经常账户的累积头寸来进行度量)。这两种方法都只取得了一定的成功,对不同货币和不同时期而言估计结果都不是特别稳健(Frankel(1982a),Rogoff(1984))。

前瞻性模型

如果将 UIP 包含到 FPMM 中,那么我们就有

$$E_ts_{t+1}-s_t=r_t-r_t^* \tag{18}$$

$$s_t=p_t-p_t^*=(m-m^*)_t-\phi(y-y^*)_t+\lambda(r-r^*)_t \tag{19}$$

将 UIP 条件中的 $r-r^*$ 代进来并进行重新排列后就得到

$$s_t=[1/(1+\lambda)]z_t+[\lambda/(1+\lambda)]E_ts_{t+1} \tag{20}$$

其中,$z_t=(m-m^*)_t-\phi(y-y^*)_t$。向前重复迭代就得到

$$s_t=[1/(1+\lambda)]\sum_{i=0}^{\infty}[\lambda/(1+\lambda)]^iE_tz_{t+i} \tag{21}$$

我们在这里施加了一个横截性条件。由(21)可知,只有当与 z_{t+i} 未来路径有关的新信息或消息出现时即期汇率才会发生变动。在 Dornbusch - SPMM 中,如果价格及其滞后值对商品市场上超额需求 X_t 有所反应,那么我们就可以在系统中引入惯性:

$$p_t - p_{t-1} = \delta X_t \tag{22}$$

当实际汇率发生贬值时(即 s_t 增加),超额需求就会增加:

$$X_t = q_1(s_t + p_t^* - p_t) \tag{23}$$

使用 UIP 条件和货币需求方程就可以得到与(21)相类似的形式,只不过汇率方程现在包含惯性:

$$s_t = \theta_1 s_{t-1} + \lambda \sum_{j=0}^{\infty} \theta_2^j E_t z_{t+j} \quad 0 < (\theta_1, \theta_2) < 1 \tag{24}$$

在这里取决于相对货币供给和实际产出。组合平衡模型(PBM)也可表示为(24)的形式,只是要注意到如下情形就可以做到这一点:我们在这里结合风险溢价就可以对 UIP 条件进行修正,风险溢价取决于国内债券 B_t 和外国债券 B_t^* 之间的相对资产持有量。

$$r_t - r_t^* = E_t s_{t+1} - s_t + f(B_t / B_t^*) \tag{25}$$

现在,在汇率方程(25)中,基础经济变量向量 z_t 包括债券的相对持有量。

基础经济变量 z_{t+j} 未来值的预测结果取决于 t 时刻的信息,因此方程(25)可以简化为一个只包含基本经济变量的纯粹后向方程(如果我们忽略任何隐含的交叉方程 RE 约束条件),实证文献经常采用这种形式来对模型进行检验。然而,如果有谁愿意为基础变量设定一组明确的 Var 预测方程,那么他就可以考察(24)中前向展开项的所有潜力,这实际上是交叉方差限制。

27.2　模型检验

正如我们在上一节中所分析的那样,对 SPMM 的检验涉及即期汇率对相对货币存量、利率等进行回归,而对 PBM 的检验则还要包括其他资产存量。如果我们忽略超通货膨胀时期,那么这些模型在双边即期汇率的变动预测中不是很成功,尤其是在使用 1945 年以后的数据时。有些模型在较短的子样本期内表现得相当好,但对整个时期而言却表现得很糟糕。Meese(1990)对这些模型的绩效进行过早期研究。他对一个一般的方程进行了估计,这个方程大体上包含了上述全部理论:

$$\begin{aligned} s_t &= a_0 + a_1(L)(m - m^*)_t + a_2(L)(y - y^*)_t + a_3(L)(r - r^*)_t \\ &\quad + a_4(L)(\pi - \pi^*)_t + a_5(L)(F - F^*)_t + \varepsilon_t \end{aligned} \tag{26}$$

其中,F = 国内居民持有外国资产的存量,F^* = 外国居民持有国内资产的存量。Meese(1990)重复 Meese 和 Rogoff(1983)的早期检验,用期数据对方程(26)进行回归,然后将其用来对 1 个月、6 个月和 12 个月期限进行样本外预测。然后加入新的数据,再重复预测和估计过程。预测所使用的是右边变量的实际未来值。然后,他将(26)的均方根预测误差与基准情形进行比较,基准是用汇率随机游走模型的"不变"预测("no change" prediction)来给出的。Meese 发现,对(26)而言,在所有情形中使用基本经济变量的预测效果都比随机游走假设下的结果更差。

Meese(1990)将基于基础经济变量的模型失败原因归结为变量的度量误差、不恰当的估计技巧甚至变量的遗漏(因为他对众多变量都进行了试验)。他认为这些模型的失败应归因于像 PPP 条件那样的基础关系的脆弱性、在货币需求函数所发现的不稳定性和关于预期(代理人的预测不会违背理性预期公理)的大量调查数据。

PPP - LOOP

(26)中的一些变量可能是 I(1)序列,因此(为了避免统计问题),最近的多数研究都对数据的协整性质进行了考察,证实了 s_t 与相对价格(PPP)之间和 s_t 与"基础经济变量"如货币供给和

产出之间的协整关系。用布雷顿森林体系后浮动汇率时期的数据对 PPP 单方程进行检验通常都会发现相对价格和名义汇率都存在协整关系（Rogoff(1996)）。众所周知，标准检验的零假设是“不存在协整关系”，对短期数据而言拒绝零假设的功效非常的低（单位根检验的功效取决于数据长度而不是数据频率，因此，使用季度数据来替代年度数据是不能增加“检验功效”的）。克服这一问题的一种方法是使用更长时期的数据，使用大约一个世纪的数据的研究确实发现了支持长期 PPP 的大量证据（如 Lothian 和 Taylor(2000)，Taylor(2001)）。

对 PPP 的其他研究使用的是面板数据。人们发现面板数据能提高协整检验的功效，很多面板检验使用的是布雷顿森林体系后的数据，结果发现了长期 PPP 的支持证据（如 Frankel 和 Rose (1996)，Taylor 和 Sarno(1998)）。解释这些面板结果的一个困难在于，在不存在协整关系的零假设遭到拒绝时，我们并不知道 PPP 是对面板中的所有国家都成立还是只对部分国家而言成立。然而，即使其他协整检验给出的是有点冲突的证据，绝大多数经济学家还是会接受 PPP 在长期中(5～10 年)成立的结论，尽管结论对 PPP 存在大而持久的偏离。

货币性基本经济变量

Mark(1995)考查了“基本经济变量”在解释汇率短期变动和长期变动方面的有用性。Mark(1995)用货币模型来决定基本经济变量 $z_t=(m-m^*)-\alpha(y-y^*)$，汇率进行缓慢的调整：

$$s_{t+k}-s_t=\delta_k+\beta_k(s_t-z_t)+\varepsilon_{t+k}$$

Mark 使用的是 1973—1991 年期间美元对加拿大元、德国马克、日元和瑞士法郎之间的季度汇率数据。上述汇率方程是误差修正模型的一种简单形式，误差修正项为 $(s-z)_t$。Mark 发现，当 k 从 1 个季度增加到 16 个季度时，上述回归中 R^2 和 β_k 的值也会增加。这与我们在长期股票收益中所注意到的现象一样。Mark 发现，对德国马克、日元和瑞士法郎而言，模型的长期限($k=16$)样本外预测效果比随机游走模型好。上述分析不是对“完全设定的”货币模型进行检验，而是为了说明“货币性基本经济变量”对长期汇率的预测是很有用的（尽管在短期内这是不必要的）。

Mark(1995)承认 β_k 的 OLS 估计值是有偏估计（Stamburgh(1999)）。为了计算偏离程度，他在如下模型中使用了蒙地卡罗模拟技术：s_t 服从随机游走，基本经济变量 z_t（如货币供给）是持续的 AR(1)过程。

$$\Delta s_t=\alpha_0+\varepsilon_{1t}$$

$$z_t=\gamma_0+\gamma_1 z_{t-1}+\varepsilon_{2t}$$

他使用参数的代表性值来生成 s_t 和 z_t 序列，然后使用“拟合数据”来分析上述误差修正模型。他发现使用拟合数据所得到的 β_k 的 OLS 估计值的平均值为正，因为根据构造 Δs_t 服从随机游走过程。这为他提供了真实数据偏差的一个度量指标（参见第 2 章）。例如，如果 MCS 中 β_k 的均值为 0.20，而实际数据所给出均值却为 0.205，那么对偏差进行调整的 β 部分为 adj $-\beta_k=0.205-0.20=0.005$。因此，实际数据中看似更大的 $\hat{\beta}_k=2.05$ 具有误导性，因为它是有偏的，偏差修正值为 0.005。这意味着在实际数据中基本经济变量对未来即期汇率的变动几乎没有什么影响。Mark 在对 β 进行偏差调整后再使用实际数据（和经验标准差）却发现，基本经济变量对长期限(>4 年)而言有一定的解释力，但对短期限(<4 年)而言却没有解释力。

协整检验适用于汇率的货币模型，其长期形式为，

$$s_t=(m-m^*)_t-\phi(y-y^*)_t+\lambda(r-r^*)_t \qquad (27)$$

UIP 条件意味着

$$(r-r^*)_t=E_t\Delta s_{t+1} \qquad (28)$$

在 RE 假设下，预测误差是 I(0)序列（否则根据 t 时刻的已知信息可以对其进行预测），从实证结果来看 s_{t+1} 是 I(1)序列，因此(28)意味着 $r-r^*$ 是 I(0)序列。在(28)中，如果 $r-r^*$ 是 I(0)序列，那么它与 s_t 之间就不存在协整关系，因为 s_t 是 I(1)序列，所以 s_t 和货币性基本经济变量之间的长期统计关系为

$$s_t=(m-m^*)_t-\phi(y-y^*)_t \qquad (29)$$

我们在这里预期$(m-m^*)$的系数为1。如果两个国家货币需求的收入弹性都为1，那么$(y-y^*)$的系数也为1。当然，可能出现的情形是，对有些国家而言，s_t、$(m-m^*)$和$(y-y^*)$都是I(0)序列，这是就不存在协整问题。

Groen(2000)及Mark和Sul(2001)使用布雷顿森林体系后多个国家的数据发现(27)中的协整关系，$\phi=1$，基于误差修正模型(ECM)的预测结果一般会比随机游走模型更好一些。显然，在Mark和Sul(2001)的面板结果中，使协整关系成立的那些国家可能也“包含”那些货币模型不成立的国家。

与之相对照的是，Rapach和Wohar(2002)使用14种汇率(对美元而言)一个世纪的年度数据来对(29)中的协整关系进行检验。使用多种方法所得出的结果却不尽一致，协整关系对法国、意大利、新西兰和西班牙而言成立，但对比利时、芬兰、葡萄牙和瑞士而言却只得到了部分支持，澳大利亚、加拿大、瑞典和英国却不支持协整关系。在协整关系成立的那些国家中，Rapach和Wohar对双变量VECM模型进行了估计：

$$\Delta s_t=\{\Delta s_{t-j},\Delta fdm_{t-j}\}+\alpha_{11}z_{t-1}+\varepsilon_{1t} \tag{30a}$$

$$\Delta fdm_t=\{\Delta s_{t-j},\Delta fdm_{t-j}\}+\alpha_{21}z_{t-1}+\varepsilon_{2t} \tag{30b}$$

“基本经济变量”是$fdm_t=(m-m^*)_t-(y-y^*)_t$和$z_t=(s-fdm)_t$。对于比利时、意大利和芬兰而言，$\alpha_{11}\neq 0$且$\alpha_{21}=0$，因此对这三个国家(而不是其他国家)而言基本经济变量是弱外生的，“基本经济变量”的非均衡驱使汇率发生变动(反之却不亦然)。事实上，这里起作用的是$(m-m^*)_{t-1}$，因为如果单独将$(y-y^*)$($I(0)$)(序列)拿进来，那么它在统计上不显著。

Rapach和Wohar(2002)使用(30a)来为Δs_t提供“样本外”预测，每年对参数更新一次，使用直到$t-1$期的数据并进行单步预测，预测只是针对z_t为I(0)序列(即s、m和y协整)的国家而进行的。他们将使用基本经济变量的预测结果与使用多种度量方式(如均方误差，MSE)的“不变”(no change)随机游走预测结果进比较后发现，对一些国家(如比利时、意大利，瑞士也有可能)而言货币模型比随机游走模型的表现更好，ECM项在Δs_t方程中是统计显著的。反之，如果ECM项在统计上不显著(如法国、葡萄牙、西班牙)，那么(30a)的预测表现就比随机游走“不变”预测要差些。也要注意的是，Rapach和Wohar(2002)的预测结果允许系数在每一个时期中都可以不同(即不是常数)，我们不知道短期“Δ”项和*ECM*项在整体预测绩效中的相对贡献。无需多言的是，$\Delta s_t=f(\Delta s_{t-j},\Delta f\mathrm{d}m_{t-j})$所给出的预测结果比简单$\Delta s_t=0$所给出的预测结果要好一些。ECM项是长期经济理论的体现，因此考察将这一项排除在外是否会导致预测绩效的大幅降低是一件非常有趣的事情。事实上，Theil U统计量$U=\mathrm{RMSE}_{fdm}/\mathrm{RMSE}_{RW}$基本上位于0.98和1.02之间，因此基础经济变量模型对随机游走模型绩效额改善是非常微小的。

总之，即使我们对货币模型进行技术上更为复杂的“再回归”，模型所给出的结果以也不会很令人满意，我们不能认为中央银行针对“适度通货膨胀”进行货币政策决策时使用的就是这个模型。

因为PPP成立，所以货币模型的缺陷似乎在于货币需求关系的稳定性上，我们已经对困难之处进行了详细的论述(例如，参见Cuthbertson(1997)及其中的其他文献)。上述“检验”假设线性成立，但可能出现的情形是基本经济变量对Δs_t的非均衡效应是非线性的。事实上，Taylor和Peel(2000)发现$z_t=(s-fdm)_t$是平稳序列，他们使用非线性ESTAR模型来对协整回归的残差进行建模，结果发现(与股票的红利价格比一样)，在“接近均衡”的地方汇率具有随机游走行为，但在远离均衡的地方汇率(围绕货币性基本经济变量)具有均值反转特征。

前瞻性模型：Var方法

在FPMM分析中，我们知道即期汇率可以用基本经济变量向前卷积的形式来进行表示：

$$s_t=(1+\lambda)^{-1}\sum_{i=0}^{\infty}[\lambda/(1+\lambda)]^i E(x_{t+i}/\Omega_t) \tag{31}$$

其中，

$$x_t = (m_t - m_t^*) - (\phi y_t - \phi^* y_t^*) \tag{32}$$

(31)两边同时减去 x_t 就得到，

$$q_t = s_t - x_t = \sum_{i=1}^{\infty} \psi^i E(\Delta x_{t+i}/\Omega_t) \tag{33}$$

其中，$\psi = [\lambda/(1+\lambda)]$，$E_t\Delta x_{t+1} = x_{t+1} - x_t$，$E_t\Delta x_{t+1} = E_t(x_{t+i} - x_{t+i-1})$ $(i > 1)$。我们可以不太严格地将 $q_t = s_t - x_t$ 称为“汇率差价”。如果(33)中的基本经济变量都是I(1)序列，那么(33)中的 Δx_{t+1} 就是 $I(0)$ 序列，因此 s_t 和 x_t 应该是协整的且 $q_t \sim \mathrm{I}(0)$。如果我们用 $z_t = (q_t, \Delta x_t)'$ 来构造一个 Var，那么很容易就可以看出(33)意味着

$$e1'z_t = \sum_{i=1}^{\infty} \psi^i e2'A^i z_t = e2'\psi A(I - \psi A)^{-1} z_t \tag{34}$$

FPMM 的这一版本所隐含的 Wald 约束条件为

$$e1' - e2'\psi A(I - \psi A)^{-1} = 0 \tag{35a}$$

或

$$e1'(I - \psi A) - e2'\Psi A = 0 \tag{35b}$$

方程(35b)是一组线性约束，表示即期汇率的 RE 预测误差取决于 t 时刻所能获得的信息，而不是由变量 x_t 来给出。我们也可以将“理论汇率差价”q_t' 定义为

$$q_t' = e2'\psi A(I - \psi A)^{-1} Z_t \tag{36}$$

然后将它与真实“差价”q_t 进行比较。为了使用 Var 方法，我们需要$\{\phi, \phi^*\}$的估计值来构造变量 $q_t = s_t - x_t$，需要 λ 的估计值来计算 ψ。FPMM 的含义是：如果所有变量都是 $I(1)$ 序列，那么

$$s_t = (m_t - m_t^*) - (\phi y_t - \phi^* y_t^*) + (\lambda r_t - \lambda^* r_t^*) \tag{37}$$

就是一个协整关系。$(\phi, \phi^*, \lambda, \lambda^*)$的估计值可以通过单方程或 Johansen 程序的协整回归来获得。然后，我们即可以构造变量 $q_t = s_t - x_t$ 和 Δx_t，并在 *Var* 中使用。

MacDonald 和 Taylor(M - T)(1993)对这一过程进行了很好的说明，他们使用的是 1976 年 1 月到 1990 年 12 月德国马克/美元汇率的月度数据。他们使用 Johansen 程序发现无法拒绝零假设：相对货币供给和相对收入的系数对于“国内”(德国)和“国外”(美国)变量而言都为 1，利率系数近似都相等但符号相反：

$$s_t = (m_t - m_t^*) - (y_t - y_t^*) + 0.049 r_t - 0.050 r_t^*$$

他们取 $\lambda = \lambda^* = 0.05$，从而 $x_t = (m - m^*)_t - (y_t - y_t^*)$。*M - T* 发现 Wald 检验果断地拒绝了 RE 交叉方程约束条件。实际“差价”q_t 的方差大于由(36)给出的理论差价 q_t' 的方差，前瞻性 FPMM 的表现很糟糕。

与 Rapach - Wohar 方法一样，*M - T* 注意到协整向量(32)的出现意味着即期汇率可以表示成动态 ECM 的形式，并使用第二阶段的回归来估计 s_t 的短期动态反应。他们所偏好的 ECM 为

$$\underset{(0.003)}{\Delta s_t = 0.0005} + \underset{(0.07)}{0.24\Delta s_{t-2}} - \underset{(0.23)}{0.42\Delta_2 m_t} - \underset{(0.34)}{0.79\Delta y_t} - \underset{(0.003)}{0.008\Delta^2 r_t^*} - \underset{(0.013)}{0.025 q_{t-1}} \tag{38}$$

其中，$R^2 = 0.14$，$SE = 32\%$，(.) = 标准差。尽管 R^2 等于 14% 意味着方程只能解释 Δs_t 变化的一个很小的部分，但方程还是通过了所有的常规性诊断检验。然后，他们进行“滚动回归”，使用(38)来对不同期限的汇率进行预测，同时对参数进行更新。他们发现，对 1 个月、2 个月、3 个月、6 个月、9 个月和 12 个月的期限而言，ECM 的均方根误差比随机游走“不变”预测要小一点点。还要注意的是，(38)中的绝大多数统计解释似乎都归因于特殊的动态项，极少的统计解释将原因归结为长期误差修正项。预测很有可能受差异项所主导，而这些项本身就接近于随机游走，因此这就是为什么所报告的均方根误差与随机游走模型近似相等的原因。*M - T* 研究是试图对 FPMM 复杂版本进行正确检验的一种变体，但对德国马克/美元即期汇率而言模型明显遭到了拒绝。

即期汇率和基本经济变量的波动性

Flood 和 Rose(1995)给出了一种检验汇率货币模型的新方法。他们将“固定汇率”(布雷顿森林体系)和“浮动汇率”时期的即期汇率波动性与“基本经济变量”的波动性进行了比较。并不令人吃惊的是,浮动汇率时期的汇率波动得更为厉害。如果货币模型是正确的模型,那么我们就能观察到从“固定”汇率向浮动汇率转变后基本经济变量(如相对货币供给)的波动性会急骤增加。对9个工业化(OECD)国家而言,Flood 和 Rose 发现,尽管与美元相关的双边汇率的条件波动率在不同汇率制度之间存在巨大差异,但没有那个基本经济变量有着较大的波动率变化。因此,我们可以合理地得出如下结论:货币模型中的经济基本变量(如货币供给、利率、通货膨胀率、产出)无法解释汇率的波动性。

在(3)中 $\phi=\phi^*$ 和 $\lambda=\lambda^*$ 并将 UIP 条件中的 $r-r^*$ 代入,我们就可以得到检验 FPMM 检验的更为正式的 Flood - Rose 方法:

$$s_t = TF_t + \lambda(s^e_{t+1} - s_t) \tag{39}$$

其中,$TF_t=(m-m^*)_t-\phi(y-y^*)_t$ 表示“传统的基本经济变量”。我们将“虚拟的基本经济变量”定义为 $VF_t=s_t-\lambda(r-r^*)_t$。在 FPMM 即方程(3)中,我们期望虚拟基本经济变量 VF_t 与传统基本经济变量 TF_t 的变化程度相同。为了能获得时间序列 TF_t 和 VF_t,唯一要求是我们必须知道货币需求函数中结构性参数 λ 和 ϕ 的代表性值。正如前面所报告的那样,Flood 和 Rose 发现,当 VF_t 的波动性在浮动汇率时期急骤增加时,传统基本经济变量 TF_t 的波动性却只有少许变动。(这一结论对 λ 和 ϕ 的合理值而言都是不变的,当我们考虑 SPMM 时该结论也同样成立。)Flood 和 Rose 认为,因为只有少数宏观经济变量的波动率出现了急骤变动,与之相伴的是汇率制度的变化,所以只依赖于基本经济变量的汇率模型都不可能完全解释现实。他们还就9个 OECD 国家连续两年的平均月度汇率方差 $\sigma^2(S)$ 与各种宏观经济变量方差之间的相关性进行了考察,结果发现 $\sigma^2(S)$ 与货币供给、利率和外汇储备、股票价格等的变化之间都不存在相关性,与产出的方差之间也只存在非常微弱的负相关关系。

理性泡沫

在前瞻性模型中出现理性泡沫是完全可能的,我们接下来简要地对这些问题进行讨论。对理性泡沫进行检验时会遇到一些严重的计量经济学困难,检验方法依赖于正确的资产收益均衡模型。拒绝无泡沫的零假设可能是因为基本经济变量的模型设定有错误。对汇率货币模型而言,这一点尤其突出,因为这些模型对即期汇率变动的统计表示相当糟糕。带有确定性泡沫项的 FPMM 给出了如下形式的即期汇率方程:

$$s_t = (1+\lambda)^{-1}\sum_{i=0}^{\infty}[\lambda/(1+\lambda)]^i E_t x_{t+i} + B_0[(1+\lambda)/\lambda]^t \tag{40}$$

其中,x_t 是货币变量集合,B_0 是在 $t=0$ 时的泡沫价值。不存在泡沫的零假设为 H_0: $B_0=0$。然而,如果 x_t 中的变量是基本经济变量的糟糕表示,那么对 B_0 的估计值可能不会等于0,因为其他未曾考虑的变量有可能有助于解释因变量。对(40)中的 $B_0=0$ 进行检验也是一个棘手的问题,因为可能出现“发散性回归”问题。然而,West(1987a)的检验却对这一问题进行了回避,并给出了任何形式泡沫(随机的或定性的)的检验方法。

Meese(1986)将 FPMM 视为基本经济变量模型来对美元/德国马克汇率(1973—1982)进行检验,得出的结论是拒绝不存在泡沫的假设。West(1987b)使用 West(1987a)中的第二类方法来进行检验,这种方法是对 Meese 的 FPMM 方法的扩展,这种方法将货币需求误差进行了考虑,货币需求误差可能会发现其他潜在的“基本经济变量”。对美元/德国马克汇率(1974—1984)而言,他发现泡沫根本不存在。由于基于基本经济变量的汇率模型的糟糕实证表现,我们只有很小的把握说理性泡沫或其他形式的泡沫是存在的。

27.3 新开放经济宏观经济学

在过去10年或更长的时间内，经济学家们致力于为宏观经济模型构筑“合理的”微观基础嵌入（Woodford（2003），Clarida、Gali 和 Gertler（1999）），这对开放经济模型产生了重大影响（Obstfeld 和 Rogoff（2000））。“严格的微观基础”的通常解释是，假设代理人在预算约束条件下追求跨期效用的最大化。我们要注意的是，这类模型也可以用 SDF 方法来进行演绎，SDF 方法试图解释风险资产的需求和股权溢价。此外，因为我们拥有显示跨期效用函数，这意味着我们可以分析各种解对整体福利的影响。

在下面的内容中，我们对这些开放经济模型进行概述。迄今为止，这类模型的实证及其在政策分析中的应用还没有得到广泛的研究，有兴趣的读者可以参考 Lane（2001）及 Sarno 和 Taylor（2002）来获取更多的信息。

与跨期消费组合模型一样，新开放经济模型（new open - economy model，NOEM）都是随机的一般均衡模型，我们通常只能对其进行校准和模拟，而无法进行估计。在 Obstfeld 和 Rogoff 文章中，“生产者 - 消费者”（producer - consumer）追求终生效用函数的最大化：

$$U = \sum_{t=1}^{\infty} \delta^t \left[\frac{\sigma}{\sigma - 1} \ln C_t^{(\sigma-1)/\sigma} + \frac{k}{1+\varepsilon} \ln \left(\frac{M_t}{P_t} \right)^{1-\varepsilon} - \frac{k}{\mu} y_t(z)^{\mu} \right]$$

动态预算约束条件为（对所有 t），

$$B_{t+1} - B_t + rB_t + \frac{M_t - M_{t-1}}{P_t} = \frac{P_t(z)}{P_t} - C_t - T_t$$

其中，C_t 是总消费，M_t 是货币供给，B_t 是全世界范围内交易的无风险债券，r 是实际利率，T_t 是税收收入，y 是实际产出。中括号中的项 $-y(z)$ 表示如下事实：高产出意味着少闲暇和低效用。参数 z 是模型中代理人个数的反映指标，它在一个连续集合中取值。

在模型中，LOOP 对价格为 $p_t(z)$ 的个体商品而言是成立的，PPP 对总价格水平而言（即 $P = SP^*$）也是成立的。模型假设实际利率平价成立，它将实际利率和名义利率联系起来了。

由于短期内价格具有粘性，因此货币供给的增加在短期内具有真实效应，但货币在长期中不是中性的，因为通过经常账户能产生财富累积。例如，在粘性价格条件下，货币供给的增加会导致名义利率的下降，通过 UIP 就可以知道汇率会贬值。后者会导致出口增加（进口减少）、产出的需求增加。垄断竞争意味着价格高于边际成本，厂商所提供的产出比不变价格时的需求高。“国内”经常账户在短期内出现盈余，这会导致更高水平的长期财富。后者意味着货币在长期内不是中性的（即使在对价格进行调整后，结果依然如此）。模型中没有汇率超调，汇率在长期内没有必要贬值。可以证明，货币扩张毫无疑问能增加福利，因为它能减弱垄断价格的扭曲效应。我们希望一般均衡模型的简短叙述能说明模型的多样性和如下可能性：模型所给出的结论可能与前面讨论的“标准”货币模型不同。

这篇文献中出现过众多核心问题，包括：是买方货币（当地货币）还是卖方货币（生产者货币定价）的价格具有粘性？工资和价格是否具有粘性（Betts 和 Devereud（2000））？Bergin（2003）对澳大利亚、加拿大和英国等“小型”开放经济的 NOEM 进行了统计检验（与校准相对照），所得出的结论却不尽一致。模型试图解释名义汇率、经常账户、价格、货币和实际利率、技术冲击、国外利率、国外需求和偏好变化。他对带约束模型和不带约束模型都进行了估计，使用似然比检验来进行比较（在进行极大似然估计之后）。尽管在所考虑的三个国家中有两个国家没有拒绝带约束的模型，但在对三个国家的汇率或经常账户进行预测时，模型却无法击败随机游走模型。价格刚性对模型来说非常重要，因为弹性价格版本的模型遭到了拒绝——但就我们所知道的 PPP 实证结果而言，这几乎不是什么重要见解。

从理论层面来看，这类动态一般均衡模型已经在多个方向进行了扩展，价格粘性机制更为

精细（如市场定价与蹒跚价格设定（stagger price setting）），偏好也划分得更为精细（如消费和闲暇的不可分离性、国内和国外商品替代程度的可变性），有的还引入了资本物品和时变风险溢价。并不令人吃惊的是，假设（包括效用函数参数）的改变会导致模型预测结果发生变化，因此，“对某个特殊的新开放经济模型达成一致意见几乎是不可能的”（Sarno 和 Taylor（2002），p. 165）。常用的语言是，“这一领域还需要进行更为深入的研究”，在现实世界的政策分析接受这类模型之前，这句话肯定是对的。

27.4　小结

对读者而言，注意到用基本经济变量来解释即期汇率变动的各种模型是非常重要的，原因在于这些模型有助于关键经济政策的制定，正如我们在公开讨论中所指出的那样。

* 通过对传统货币模型进行微小改变来提高模型的统计绩效似乎是不大可能的事情。就目前而言，似乎可能的情形是，各种模型的规范检验结果是拒绝模型本身。

* 在短期内，比如说一年，货币性基本经济变量一般无助于即期汇率变动的预测。只有在四年以上的更长时期中，基本经济变量才对某些货币（但不是很多）具有预测功效（Mark（1995），Rapach 和 Wohar（2002）），两者之间的关系可能是非线性关系（Sarno 和 Peel（2000），Taylor、Peel 和 Sarno（2001））。

* 我们必须承认，对通货膨胀适度的工业化国家而言，我们仍不能确定汇率的基本决定因素。但这些模型背后的概念和思想（如 PPP、UIP 和相对货币供给）对政策制定者而言仍然具有重要的指导作用——尽管政策制定者在制定汇率时会考虑本章所讨论的模型和其他因素。

* 新开放经济模型给出了在不完全竞争条件下的跨期一般均衡模型中分析货币政策的机制，模型还可以对各种政策进行福利分析。这是一个有趣的进展，但到目前为止，对绝大多数合适的模型和实证效果而言，人们很难达成一致意见。因此，政策性结论仍然具有很强的尝试性。

* 因为这一章和前面几章中所讨论的“结构化”货币模型存在统计检验方面的不充分性，所以汇率模型缺乏清晰的政策含义。使用类比方法并注意到，就汇率行为的解释而言，不是“皇帝没有穿衣服”，而是因为他有太多的衣服却没有一件看上去非常合适。这就致使政策制定者试图通过中央银行的协调干预或构建货币区域或结成像欧元区那样的共同货币联盟等方式来减缓实际汇率大幅波动的严重性。

第28章

市场风险

目的

* 对在险价值(Var)进行定义,说明单一资产和投资组合的Var如何进行度量,投资组合的收益在这里是个体资产收益的线性函数,资产收益服从(多维)正态分布。这种方法就是著名的Delta-Normal方法或方差-协方差(VCV)方法。

* 概述在使用Delta和Delta+Gamma近似技巧来计算期权头寸的Var时所存在的局限性。

* 说明“映射”如何减少Var计算中的输入数量,说明我们如何将一些非线性投资工具(如债券)“映射”到近似线性框架之中,从而可以继续使用VCV方法。

* 考察估计投资组合Var的非参数方法,如历史模拟法,说明如何使用自助法来对其进行扩展。

* 说明如何使用Delta方法(即线性近似法)、Delta-Gamma近似法和蒙特卡罗模拟方法来对包含期权的投资组合的Var进行计算。

* 分析传统Var方法的一些变体,如压力测试法、更坏情形的情景分析法和“极值理论方法。

在这一章里,我们讨论市场风险,市场风险是由可交易资产如股票、债券、外汇、期货和期权的价格变化所导致的风险。很多金融机构(及一些大的工业公司和商业公司)持有多种资产的净头寸。处于谨慎方面的原因,金融中介需要计算其资产组合的整体“美元”市场风险,我们通常称其为在险价值(Var)。此外,管理当局使用Var来设定能控制市场风险的最低资本需求。在Var方法的发展过程中,投资银行J. P. 摩根起了重要作用,主要内容以Risk Metrice™为题进行了正式出版。如果你对期权了解不多,那么你可以参考Cuthbertson和Nitzsche(2001b)。

28.1 Var的度量

什么是Var? 如果在5%的临界值下每天的Var是1亿美元,这意味着金融机构在20天里只有1天的损失会大于1亿美元。值得注意的是Var可以用以下几种方法来进行计算:方差-协方差方法(the variance - covariance method)、历史模拟自助法(historical simulation - bootstrapping techniques)、蒙特卡罗模拟(Monte Carlo simulation, MCS)、压力测试(stress testing)和极值理论(extreme value theory)。很多具体讨论都是基于J. P. 摩根Risk Metrice™(1996)中所陈述的风险度量方法。

个体资产的风险可以用收益的概率分布来刻画。尽管人们经常使用正态分布的标准差,但对一个特殊分布而言,还没有一个令人满意的风险度量指标。对收益独立同正态分布(niid)的

个体资产而言，我们可以用方差（或标准差）来度量风险。对收益服从正态分布的那些资产来说，我们有90%的把握说真实收益等于预期收益加或减1.65σ，σ在这里表示收益的标准差。换句话说，我们预期真实收益小于$\mu - 1.65\sigma$或大于$\mu + 1.65\sigma$的概率仅5%。

更为深入地说，如果假设平均收益为零，那么我们可以将1.65σ视为“下跌风险”（downside risk）的一个度量指标（在5%的概率下）。因此，如果我们在某种资产上的净头寸为V_0美元，那么下跌Var值为：

$$\text{Var} = V_0(1.65\sigma)\text{美元} \tag{1}$$

假设资产组合的Var在一天内为1亿美元，那么这个Var可以等价地表示为：

在20天的时期中，19天内的预期最大损失为1亿美元，或者说只在某1天的预期损失会超过1亿美元。

如果资产收益不服从正态分布，那么90%的置信区间就不是±1.65σ。然而，正如我们即将看到的那样，对许多资产收益而言，正态分布假设是一种合理的近似。

Var：资产组合

对许多资产组合而言，假设投资组合的价值改变量dV_p与资产收益之间存在线性关系是合理的：

$$dV_p \sum_{i=1}^{n} N_{0,i}dP_i = \sum_{i=1}^{n} V_{0,i}R_i = \sum_{i=1}^{n} V_{0,i}(dP_i/P_i) \tag{2}$$

其中，$N_{0,i}$表示资产i的数量，$V_{0,i} = N_{0,i}P_{0,i}$表示资产i的（当前）持有价值，R_i是资产i的收益。对权益资产而言上式精确成立，对其他资产（如付息债券、FRA*s*、FRN*s*、远期、期货和互换）也是一种合理的近似。由方程（2）可以得出投资组合的标准差：

$$\sigma_{dV_p} = \left[\sum_{i=1}^{n} V_{0,i}^2\sigma_i^2 + \sum_{i\neq j}\sum V_{0,i}V_{0,j}\rho_{ij},\sigma_i\sigma_j\right]^{1/2} \tag{3}$$

方程（3）可用来计算投资组合的标准差，根本不用考虑资产收益的分布形式。但标准差本身却无法使我们确定5%的下限值，除非我们假设资产收益服从某个特殊的分布。特别地，如果我们增加资产收益的（多维）正态分布假设，那么5%的分散风险值（或分散Var，Diversified Var）为$\text{Var}_p = 1.65\sigma_{dV_p}$，我们可以将其写成一种更为紧凑的形式：

$$\textbf{分散 Var}_p = \sqrt{ZCZ'} \tag{4}$$

其中，$Z = [\text{Var}_1, \text{Var}_2, \cdots, \text{Var}_n]$，每种资产的Var分别取值为$\text{Var}_i = V_{0,i}(1.65\sigma_i)$，$C$是相关系数矩阵。例如，对$n=3$时的相关系数矩阵为：

$$C = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix}$$

线性和正态分布假设在这里非常关键。资产价值的变化有时与线性完全一致（如股票），即使两者不一致，我们也经常使用逼近技巧（进行一阶Taylor级数展开）来施加线性性。注意，如果组合价值dV_p与资产收益之间的关系高度非线性（如股票期权），那么即使标的资产（如股票）的收益服从正态分布，资产组合价值的分布也不会服从正态分布。因此，对由期权所组成的一个投资组合而言（举个例子），完整分布需要准确计算Var（对给定的任意百分位数而言），而不能简单地使用“1.65σ”规则。

假设我们可以使用线性和（多维）正态分布假设，为了计算分散Var，我们需要相关系数矩阵、每种资产的（目前）持有量$V_{0,i}$和收益的波动率。这就是众所周知的“方差－协方差”（VCV）方法。

当所有资产都为多方头寸（即$V_{0,i} > 0$）且所有资产收益之间完全正相关（从而为单位矩阵，所有元素都为1）时，Var会出现更坏的情形（worse case）。因此，由（4）有：

$$\textbf{更坏情形中的 Var} = \textbf{Var}_1 + \textbf{Var}_2 + \cdots + \textbf{Var}_n$$

表 1 说明了如何很容易地用矩阵形式来表述上述公式。三种资产(股票)的持有量都为 10 000 美元,但资产 2 是卖空头寸。所有的相关系数都为正,且值很大。“更坏情形”或未分散 Var (undiversified Var)是 1996 美元,但分散 Var 要小得多,只有 783 美元。这是因为,尽管资产 2 与其他资产之间具有正的相关系数,但它是卖空头寸,这会在资产组合中产生负的相关性。

表 1　　由 3 种资产构成的投资组合的 Var

资产	资产价值	标准差	Var	Var 的绝对值		相关系数矩阵		
1	10000	5.41%	894	894		1	0.962	0.403
2	-10000	3.04%	-502	502		0.962	1	0.610
3	10000	3.64%	600	600		0.403	0.610	1
					个体资产的 Var	894	-502	600
	更坏情形中的 Var			1996	分散 Var	783		

波动率的预测

VCV 方法是一种参数化方法,因为我们需要对波动率和相关系数进行预测。许多实践者(如 Risk Metrics™)都使用指数加权移动平均方法(EWMA)来预测每日波动率:

$$\sigma^2_{t+1|t} = \lambda\sigma^2_{t|t-1} + (1-\lambda)R_t^2 \tag{5}$$

R_t^2 在这里是每日收益的平方,参数 $\lambda = 0.94$,它能使历史预测误差达到最小(对给定的期限(比如说 100 天)和资产类别而言)。类似地,协方差的预测值为:

$$\sigma^2_{xy,t+1|t} = \lambda\sigma^2_{xy,t|t-1} + (1-\lambda)X_tY_t \tag{6}$$

X_t 和 Y_t 在这里是两种资产 X 和 Y 的每日收益。因此,相关系数的预测值为:

$$\rho_{xy} = \sigma_{xy}/\sigma_x\sigma_y \tag{7}$$

上述等式都是递归式等式,当 X_t 和 Y_t 的新数据到达后,我们每天都可以对这些等式进行更新。

更长的期限

对比一天还要长的时期而言,我们如何预测波动率呢?首先,我们要在给定的时期中重新计算收益的标准差(直接法)。例如,我们可能使用过去 36 个月(比方说)的月度股票收益数据来计算方程(5)中的月度标准差。一种广泛使用的方法是“$\sqrt{T}$法则”,这种方法基于如下事实:如果每日价格变化(的对数值)独立同分布,那么

$$\sigma_T = \sqrt{T}\sigma \tag{8}$$

这里的 σ 是每日标准差的预测值(我们将时间下标省略了),T 是预测期限中的交易天数。例如,在包含 25 个交易日的时期中(接近于一个月),标准差为 $\sigma_{25} = \sqrt{25}\sigma = 5\sigma$。直接法要求人们在所选择的任何“新”时期中对所有资产的方差进行重新计算,因此平方根法则的计算工作量要小得多(对一个月的期限而言,波动率的计算精度很高)。

Delta 和全局评价法

Black - Scholes(1973)将期权定价公式用来对欧式期权(即无提前执行)进行定价。我们只考虑股票期权,期权合约中的标的资产是某家公司的股票(不支付红利)。假设股票价格服从对数正态分布过程(使用随机微积分)就可以导出期权定价公式(参见 Cuthbertson 和 Nitzsche (2001b))。欧式看涨期权的 Black - Scholes 公式看上去很吓人:

$$C = SN(d_1) - SN(d_2)PV = SN(d_1) - SN(d_2)Ke^{-rT} \tag{9}$$

其中,

$$d_1=\frac{\ln(S/PV)}{\sigma\sqrt{T}}+\frac{\sigma\sqrt{T}}{2}=\frac{\ln(S/K)+(r+\sigma^2/2)T}{\sigma\sqrt{T}},$$

$$d_2=d_1-\sigma\sqrt{T}=\frac{\ln(S/K)+(r-\sigma^2/2)T}{\sigma\sqrt{T}},$$

C 是看涨期权的价格(赎回溢价),

r 是期限为 T 的无风险(连续复利)利率,

S 是当前股票价格,

T 是到期时间(以年为单位),

PV 是执行价格 K 的现值($=Ke^{-rT}$),

σ 是股票(连续复利)收益的年度标准差,

$N(\cdot)$ 是累积正态分布。

看跌期权的定价公式与之类似。就我们的目的而言需要关注的是,看涨(看跌)期权溢价是标的股票价格 S 的非线性函数(见图 1)。

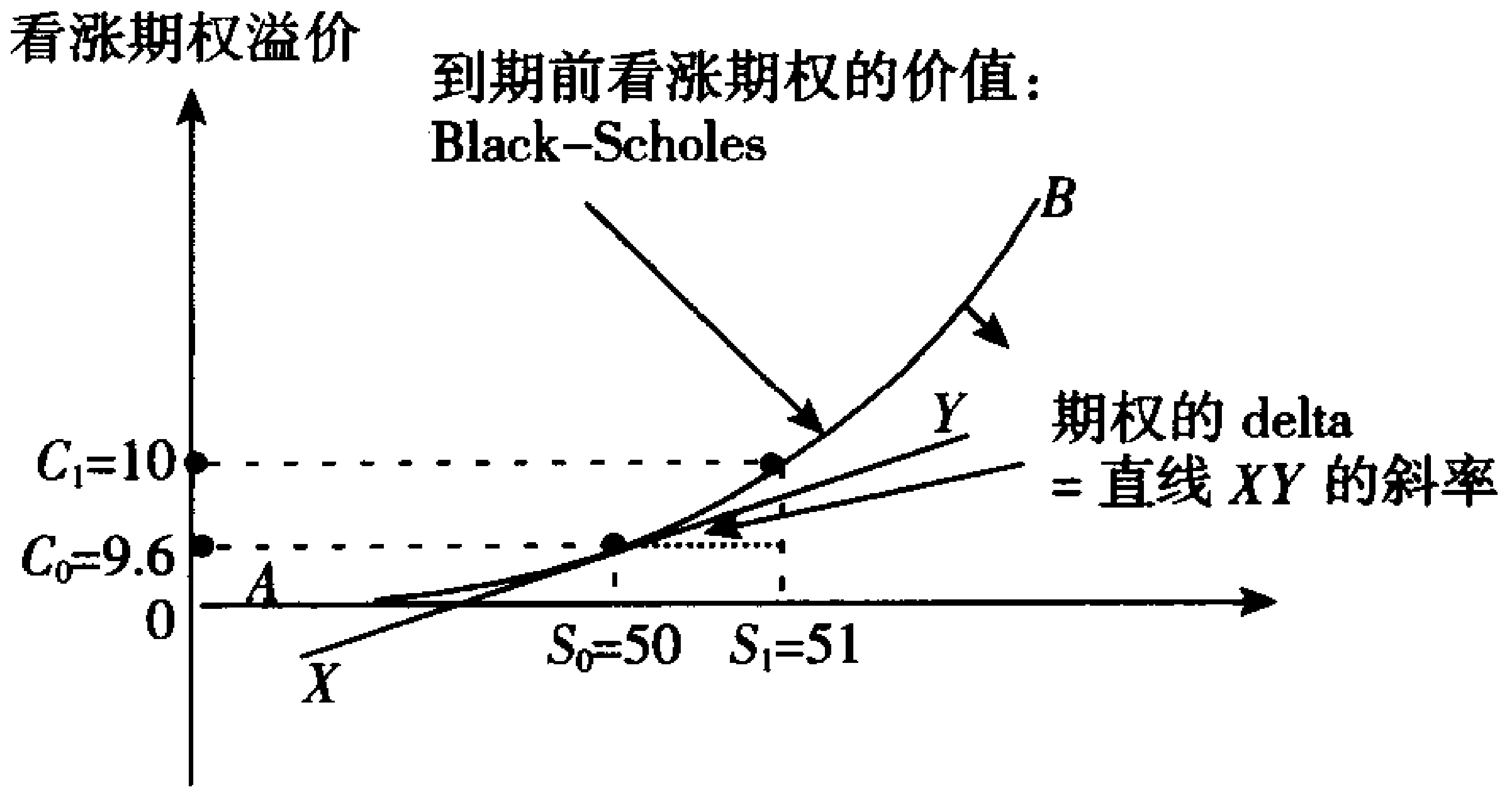


图 1　Black - Scholes 期权价格:看涨期权多头

对看涨期权而言,赎回溢价是 S 的递增函数,看涨期权的"Delta"定义为:

$$\Delta_c=\partial C/\partial S=N(d_1)>0$$

从图 1 可以看出,当 S 从 50 增加到 51 时,期权价格的实际变化是 0.4。期权的 Delta 就是直线 XY 的斜率,这为期权价格实际变化提供了一阶近似值(即 $dC=\Delta_c dS$),但也低估了实际的变化值。然而,如果我们愿意使用一阶近似值,那么 $dC/C_0=(S_0/C_0)\Delta_c(\mathrm{d}S/S_0)$,从而期权收益是股票收益 $\mathrm{d}S/S_0$ 的线性函数,因此,我们可以用方差 - 协方差方法来计算 Var。当然,如果 S 的变化很大,那么这种近似效果就会很差,这也是我们为什么还要用其他方法如蒙特卡罗模拟(MCS)来获得投资组合(包含期权)更为精确 Var 的原因。

对看跌期权而言,溢价 P 与 S 之间是一种负相关关系:

$$\Delta_p=\partial P/\partial S=N(d_1)-1<0$$

观察图 1 就会注意到,看涨期权的 Delta 值在 0(对小的 S 值而言,也就是说,在期权处于"虚值状态"的时刻)和 1(对高的 S 值而言,也就是说,在期权处于"实质状态"的时刻)之间变化。因此,Delta 的值随着 S 的变化而变化。这就是众所周知的期权 Gamma(见图 2),Gamma 值对看涨期权和看跌期权都相同:

$$\Gamma(\text{看涨期权或看跌期权})=\frac{\partial\Delta}{\partial S}=\frac{\partial^2 f}{\partial S^2}=\frac{N'(d_1)}{S\sigma\sqrt{T}}\geqslant 0 \qquad (10)$$

其中,$N'(x)=\dfrac{e^{-x^2/2}}{\sqrt{2\pi}}\geqslant 0$,$f=C$ 或 P。期权 Delta 和 Gamma 的显式公式可以由 Black - Scholes 公式直接导出。"Gamma"表示曲率的变化(也就是在不同 S 值处的 Delta 变化)。显然,无论看涨期权是深度实值期权($S\geqslant K$)还是深度虚值期权($S\leqslant K$),Γ 都是非常稳定的,但对执行价格与市场

价格很接近的期权而言它的变化非常大（见图 2）。为了得到负的 Gamma 值，投资者必须持有看涨期权的空头或持有看跌期权。

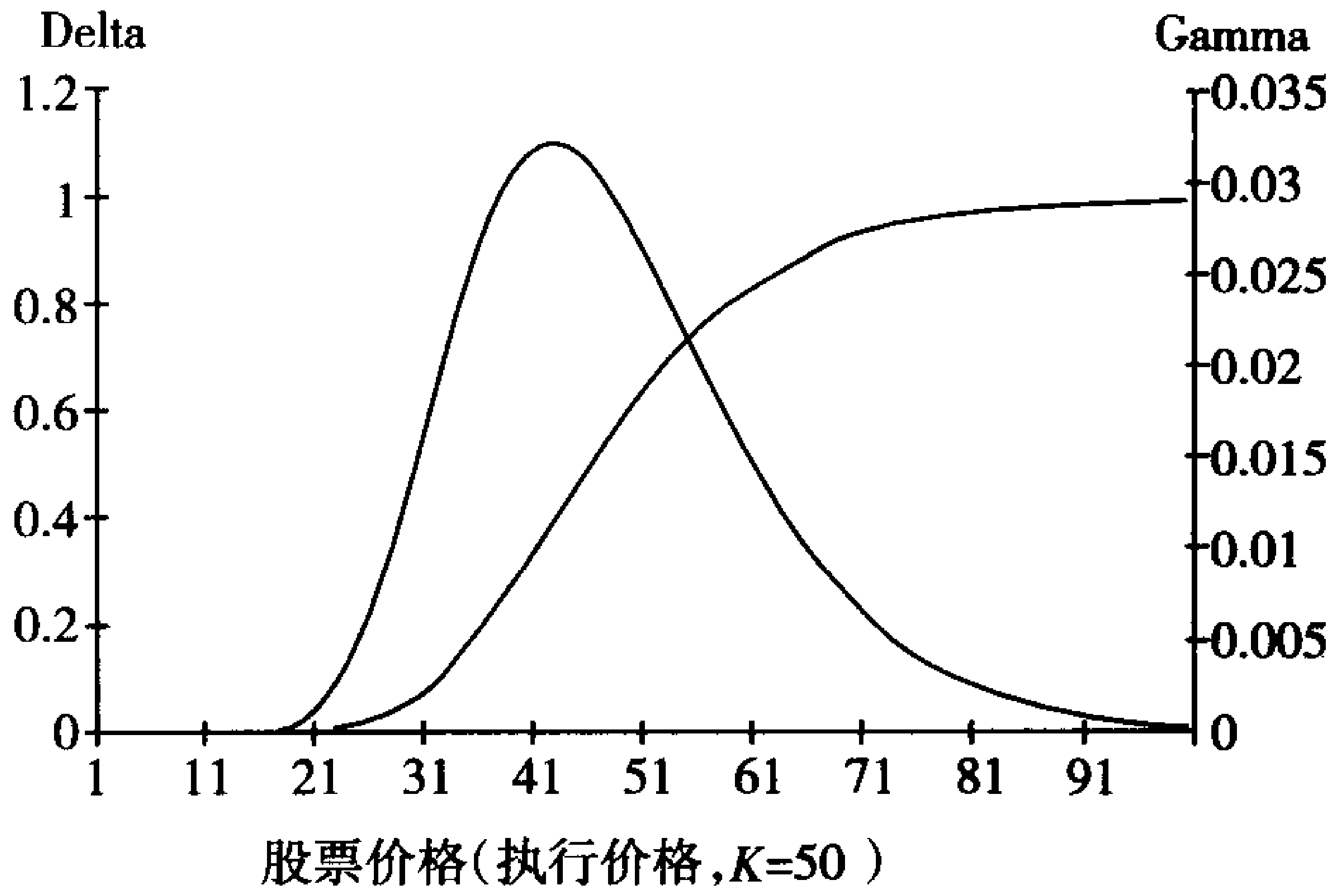


图 2　Delta 和 Gamma：看涨期权多头

在投资组合的价值非线性地依赖于资产收益的情形中（如期权），关键问题在于，投资者是使用线性（或二阶）近似值还是使用“全局”评价法。第二个关键问题是标的资产的收益是否服从（多维）正态分布。

为了说明如何使用方差 - 协方差方法来计算“非线性”资产组合的 Var，考虑如下情形。令 N_i 为 i 类资产的数量，资产的价值为 V_i。V_i 可能是期权的价值，此时 V_i 与标的资产的价格（股票价格）S 之间是一种非线性关系：

$$V_p = \sum_{i=1}^{n} N_i V_i \tag{11}$$

如果每种资产的价值 V_i 都取决于标的资产的价格 S_i（例如，期权的价值依赖于股票价格），那么投资组合价值变化的一阶近似值为

$$\mathrm{d}V_p = \sum_{i=1}^{n} N_i S_i(\partial V_i/\partial S_i)(\mathrm{d}S_i/S_i) \equiv \sum_{i=1}^{n}(V_0,i\Delta_i)R_i \tag{12}$$

其中，$V_{0,i} \equiv N_i S_i$，$\partial V_i/\partial S_i = \Delta_i$，标的资产的收益为 $R_i \equiv \mathrm{d}S_i/S_i$。这就是众所周知的 Delta 定价方法，因为一阶导数就是指资产（如期权）的 delta。例如，如果是期权溢价，那么 $\partial V_i/\partial S_i = \Delta_i$ 就是期权的 *Delta*。如果附加资产收益 dS_i/S_i 服从（多维）正态分布的假设，那么我们就可以用（标准）正态分布的分位数来确定 Var。这就是 delta - normal 方法术语的由来。注意，delta - normal 方法需要期权合约中标的资产收益的方差和协方差的预测值。如果我们将单个期权头寸的 Var 定义为 $\mathrm{Var}_i \equiv 1.65(V_{0,i}\Delta_i)\sigma_i$，$\sigma_i$ 是股票收益的标准差，那么方程（4）用通常的方法就可以得到分散 *Var*。（也要注意的是，（12）是（2）的一般化，因为如果 V_i 是股票价格，那么 $\partial V_i/\partial S_i = 1$，$N_i S_i = V_{0,i}$ 就是投资于股票的初始财富。）

然而，delta - normal 方法对期权（特别是实值期权）而言可能非常不准确，因为期权的 delta 只是非线性价格反应的一阶近似。因此，对期权而言，我们能刻画得更为深入一些，用二阶（Taylor级数）近似值来逼近期权组合的价值变化：

$$\mathrm{d}V_p = \sum_{i=1}^{n} N_i S_i \Delta_i(\mathrm{d}S_i/S_i) + (1/2)\Gamma_i S_i^2(\mathrm{d}S_i/S_i)^2 \tag{13}$$

Γ_i 在这里是单个期权的 Gamma 值。使用上述公式来计算价值的变化就是所谓的 Delta - Gamma 方法。然而，一个问题是，$(\mathrm{d}S_i/S_i)^2$ 项的出现意味着价值变化不再是收益的线性函数，因此我们无法使用正态分布已知的“截断点”（例如，对 5% 的右尾而言断点值为 - 1.65）。正如我们即将看到的那样，方程（13）可以用来计算期权组合的价值（近似）变化，这里的股票收益用 MCS 来生成。

我们使用方差 - 协方差方法来对非线性偿付（如期权）进行更为深入的逼近。为了便于说明问题，考虑某种期权，其 Black - Scholes 价格为 $P = f(S, r, y, \sigma, T - t)$。二阶 Taylor 级数展开式为

$$dP = df = \frac{\partial f}{\partial S}dS + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}(dS)^2 = \Delta S\frac{dS}{S} + \frac{1}{2}\Gamma S^2\left[\frac{dS}{S}\right]^2 \tag{14}$$

为了正确地应用 Delta - Gamma 近似技巧，我们需要拟合一个与 df 相匹配的分布。假设 dS/S 服从正态分布，均值为 0，标准差为 σ，可以证明 df 的前三阶矩为：

$$\begin{aligned} \mu_{df} &\equiv E(df) = (1/2)S^2\Gamma\sigma^2 \\ E(df^2) &= S^2\Delta^2\sigma^2 + (3/4)S^4\Gamma^2\sigma^4 \\ E(df^3) &= (9/2)S^4\Delta^2\Gamma^2\sigma^4 + (15/8)S^6\Gamma^3\sigma^6 \end{aligned} \tag{15}$$

一旦知道了右边变量 Δ、Γ（根据 Black - Scholes 公式）、当前的股票价格 S 和 σ 估计值（如隐含波动率或波动率的 EWMA 或 GARCH 度量）等的值，我们就能计算出 μ_{df}、$E(df^2)$ 和 $E(df^3)$。如果我们忽略 $E(df^3)$，那么前两阶矩就可以用来拟合正态分布的均值 μ_{df} 和方差 $\sigma_{df}^2 = E(df^2) - [E(df)]^2$，然后马上就可以知道第 5 分位数处的 Var 为 $1.65\sigma_{df}$。然而，忽略 $E(df^3)$ 在某些时候是非常危险的。就这里而言，我们可以使用一种程序来获得前三阶矩，得到（未知）分布百分位数的显性公式，这就是 Cornish - Fisher 扩展式（参加 Cuthbertson 和 Nitzsche(2001b)）。这种方法进行扩展后就可以用于期权组合。

但我们还有一些问题，因为期权组合的价值取决于其他变量的变化，在 Black - Scholes 假设下，以分红股票（或者外币资产，或期货期权）为标的资产的欧式看涨和看跌期权的溢价为 $P = f(S, r, y, \sigma, T - t)$，$P$ 在这里表示期权溢价。因此，Taylor 级数展开式为

$$\begin{aligned} dP &= df \\ &= \frac{\partial f}{\partial S}dS + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}(dS)^2 + \frac{\partial f}{\partial \sigma}d\sigma + \frac{\partial f}{\partial r}dr + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial t}dt \\ &= \Delta dS + \frac{1}{2}\Gamma(dS)^2 + \Lambda d\sigma + \rho_r dr + \rho_y dy + \theta dt \end{aligned} \tag{16}$$

在多维正态分布假设下，用 r、y 和 σ 的其他线性项来计算 Var 不存在什么问题。然而，多维正态分布假设，尤其是对 r 和 σ 而言是不现实的。因此，如果我们想用波动率的变化来计算期权的 Var 或需要更为准确的 Var（例如，在标的资产 S 的其他分布（相对于正态分布而言）假设下），那么我们应避免 Delta - Gamma 方法而使用其他方法，如蒙特卡罗模拟，这些方法可以使用我们从期权溢价结果的完整分布中得到 Var（参见下文）。

因此，如果资产或资产组合的价值与资产收益之间的非线性程度很高，那么我们不得不使用全局评价法：

价值变化 = “新”价格下的价值 - 初始价格下的价值

一般来说，在计算期权头寸的 Var 时，全局评价法常和蒙特卡罗模拟方法经常一起使用。

28.2　资产映射：计算的简化

国内的权益资产

为了计算某个国家股票的 Var，我们需要计算 n 只股票的方差和 $n(n-1)/2$ 个相关系数，计算任务相当繁重。为了简化计算，我们可以使用单一指数模型（SIM），这个模型使得资产收益的所有方差和协方差都包含在 n 个 β 中。因此，我们可以将个体资产收益与市场收益“对应”起来。SIM 假设资产 i 的收益由市场收益唯一决定（参见附录 II）。

$$R_i = \alpha_i + \beta_i R_m + \varepsilon_i \tag{17}$$

其中，R_i 是资产 i 的收益，R_m 是市场收益，β_i 是股票 i 的 β 值，假设 ε_i 为 *niid* 序列。另外，我们有如下重要假设：同期误差项不相关，$E(\varepsilon_i\varepsilon_j)=0$。（$R_m$ 也与 ε_i 独立独立。）使用“风险度量服务”数据或将 R_i 对 R_m 进行回归（比方说，使用最近六个月的每日数据）就可以得到 β_i 的估计值。包含 n 种资产的投资组合收益（权重为 w_i）为

$$R_p=\sum_{i=1}^{n} w_iR_i=\alpha_p+\beta_pR_m+\sum w_i\varepsilon_i \tag{18}$$

其中，$\alpha_p=\sum w_i\alpha_i$，$\beta_p=\sum w_i\beta_i$ 是权益组合的 β 值。因此，股票组合的收益线性地依赖于市场收益和投资组合的 β_p。可以证明，权益组合的标准差为

$$\sigma_p=\beta_p\sigma_m \tag{19}$$

持有投资组合可以将每只股票的特殊风险 ε_i 分散掉。因此，$\sum w_i^2\sigma_{\varepsilon_i}^2$ 这一项非常小，可以忽略不计。为了计算 σ_p，我们只需 n 种资产 β_i 的估计值和 σ_m 的估计值。因此，我们也无需估计 n 个 σ_i 值，更重要的是不用计算 $n(n-1)/2$ 个 ρ_{ij} 值。我们要做的是建立股票收益与（市场）指数收益之间的“映射”关系。这大大地减少了投资组合 Var 的计算量：

$$\mathrm{Var}_p=V_{0,p}1.65\sigma_p=V_{0,p}\beta_p(1.65\sigma_m) \tag{20}$$

$V_{0,p}$ 在这里表示投资组合总的美元数。注意，$w_i=(V_{0,i}/V_{0,p})$，上述方程可以表示为矩阵的形式：

$$\text{分散 } \mathrm{Var}_p=\sqrt{ZCZ'}=\sum_{i=1}^{n}\mathrm{Var}_i \tag{21}$$

$\mathrm{Var}_i=V_{0,i}(\beta_i 1.65\sigma_i)$，$Z=[\mathrm{Var}_1,\mathrm{Var}_2,\cdots,\mathrm{Var}_n]$，对 SIM 而言相关系数矩阵 C 为单位阵。这是因为，在 SIM 中个体资产收益的系统运动的唯一来源是 R_m，因为 $E(\varepsilon_{it},\varepsilon_{jt})=0$，所以所有资产的收益都同升或同降（尽管幅度不同）。

SIM 是单因子模型，唯一的因子就是 R_m。然而，如果假设有几个因子（如 Fama - French 的三因子模型）对 n 种资产的收益有影响，那么我们就可以在 *Var*“映射”框架下对这种方法进行扩展。于是，n 种资产收益之间相关系数只取决于这些基础性“风险因子”（及其 β）之间的相关系数矩阵。

Var：国外的权益资产

如果所持有的股票是其他国家的股票，那么我们应该如何计算资产组合的 Var 呢？假设 SIM 在任何单个国家内都成立，然后将通常的方差 - 协方差方法应用于国外收益（本国货币表示）和汇率上就可以解决这个问题。假设一位美国投资者持有一个德国证券组合，SIM 所给出的标准差为

$$\sigma_G=\beta_p\sigma_{DAX} \tag{22}$$

其中，σ_{DAX} 是 DAX（德国股票市场指数）的标准差，β_p 是德国投资组合的 beta。显然，这个投资组合受到 DAX 变动和美元 - 欧元汇率变动的外汇汇率风险的影响。如果投资组合的欧元价值为 $V_{0,\text{欧元}}$，那么其美元价值就为 $V_{0,\text{美元}}=S_0V_{0,\text{欧元}}$，$S_0$ 是当前的汇率（美元/欧元）。投资组合的美元价值变化（近似线性地）等于

$$\Delta V_p=V_{0,\text{美元}}(R_G+R_{FX}) \tag{23}$$

其中，

$V_{0,\text{美元}}$ 是德国股票组合的初始美元价值；

R_G 是德国股票组合的收益；

R_{FX} 是美元 - 欧元汇率的改变量（用美元/欧元来表示）。

因此，投资组合的方差取决于 R_G 和 R_{FX} 之间的相关系数 ρ（及两者的方差），用美元表示的 Var 为

$$\mathrm{Var}_p=(1.65)\sigma_dV_p=V_{0,\text{美元}}(1.65)(\sigma_G^2+\sigma_{FX}^2+2\rho\sigma_G\sigma_{FX})^{1/2} \tag{24}$$

这再一次可以用矩阵来进行表示：

$$\mathrm{Var}_p=\sqrt{ZCZ'} \tag{25}$$

其中，$Z=[V_{0,美元}(1.65)\sigma_G, V_{0,美元}(1.65)\sigma_{FX}]$，$C$ 是(2×2)相关系数矩阵，$\rho_{12}=\rho_{G,FX}$。我们可以将这种方法扩展到包括国内资产和国外资产的(两种资产构成的)投资组合情形。例如，如果美国投资者所持有的美国权益资产的价值为 $V^{US}_{0,美元}$，所持有的德国权益资产的价值 $V^{G}_{0,美元}$，那么向量 Z 就等于

$$Z=[V^{US}_{0,美元}(1.65)\sigma_{US}, V^{G}_{0,美元}(1.65)\sigma_G, V^{g}_{0,美元}(1.65)\sigma_{FX}] \tag{26}$$

变量 σ_{US} 是美国权益头寸的波动率，通过 $\sigma_{US}=\beta^{US}_p\sigma_{sp}$ 就可以对其进行计算，这里的 β^{US}_p 是美国权益组合的 Beta，σ_{sp} 是 S&P 权益指数的波动率。在这种情形中，(3×3)相关系数矩阵为：

$$C=\begin{bmatrix} 1 & \rho_{US,G} & \rho_{US,FX} \\ \rho_{US,G} & 1 & \rho_{G,FX} \\ \rho_{US,FX} & \rho_{FX,G} & 1 \end{bmatrix}$$

使用(26)就可以将上述问题在“只有国内资产”的框架中加以解决，因为个体资产的 $\mathrm{Var}_i=V_{0,i}(1.65)\sigma_i$，$Z=[\mathrm{Var}_1, \mathrm{Var}_2, \cdots, \mathrm{Var}_n]$，但对一个国外投资组合而言，还存在一个相应的双边汇率 σ_{FX} 问题。

如果没有对国外投资组合进行套期保值，那么当即期汇率与不同国家的股票市场收益之间具有低(或负)的相关关系时，风险降低仍然是可能的(从而 *Var* 要更小一些)。如果投资者对外汇头寸进行了套期保值，那么即期汇率的波动率和相关系数显然不会参与到 Var 的计算中去。因此，并不总是出现的情形是：与不进行套期保值的情形相比，对外币资产进行套期保值能产生更小的资产组合波动率。

债券

债券的价格和收益是非线性相关的。对零息债券而言，

$$P_{it}=M_i e^{-y_i t_i}$$

其中，P_{it} 表示 t 时刻的债券价格，M_i 表示 t_i 时刻债券的息票支付，y_i 是 t 时刻的(连续复利)即期收益。一阶近似结果为

$$\frac{\mathrm{d}P_i}{P_i}=-t_i\mathrm{d}y_i \tag{27a}$$

且

$$\sigma_i(\mathrm{d}P_i/P_i)=t_i\sigma(\mathrm{d}y_i) \tag{27b}$$

假设我们在时刻 t_i 从 k 种债券上获得的息票总额为 $TC_i=\sum_{j=1}^{k}C_j$。一组付息债券等价于一组零息债券，付息债券组合的价值为(t_i 时刻的息票总额为 TC_i)

$$V_{p,t}=\sum_{i=1}^{n}TC_i e^{-y_i t_i}=\sum_{i=1}^{n}V_{0,i} \tag{28a}$$

$V_{0,i}=TC_i e^{-y_i t_i}$ 是 t_i 时刻息票总额的现值(价格)。因此，

$$\mathrm{d}V_{p,t}=\sum_{i=1}^{n}V_{0,i}(-t_i\mathrm{d}y_i)=\sum V_{0,i}(\mathrm{d}P_i/P_i) \tag{28b}$$

债券组合的价值变化是与零息债券价格变化的线性函数。如果给定 $\sigma(\mathrm{d}y_i)$(EWMA 预测值)，那么由(27b)就可以计算出债券收益 $\sigma_i(\mathrm{d}P/P)$ 的标准差。此外，如果假设收益变化服从正态分布且久期可以用近似方法进行计算，那么我们就可以按照通常的方式用(28b)和 VCV 方法来计算债券组合的 *Var*。我们就将债券收益与价格之间的非线性关系“映射”到一个线性框架之中(更为深入的讨论可以参见 Cutherbertson 和 Nitzsche(2001b))。债券组合的 Var 为

$$\mathrm{Var}_p=1.65\sigma(dV_p)=\sqrt{ZCZ'}$$

其中，$Z=\{\mathrm{Var}_1, \mathrm{Var}_2, \cdots, \mathrm{Var}_n\}$，$\mathrm{Var}_i=1.65V_{0,i}\sigma_i$，零息债券价格的标准差为 $\mathrm{d}\sigma_i=\sigma(\mathrm{d}P_i/P_i)$，$C$ 是零息债券收益($\mathrm{d}P_i/P_i$)之间的相关系数矩阵。

28.3 非参数方法

上一节估计方差和相关系数的方法用一个特殊的参数化模型（如 EWMA、ARCH 和 GARCH）来刻画收益的波动率，在使用“Delta 方法”计算资产组合的 Var 时，我们通常要借助于正态分布假设。如果收益不服从正态分布，那么这会导致有偏估计或有偏预测。如果不打算为收益假设一个特殊的分布，那么我们可以使用非参数方法来估计 Var，非常流行的一种方法是历史模拟法。由种资产组成的投资组合在 t 时刻的实际每日收益（损失）为 ΔV_{pt}：

$$\Delta V_{pt} = \sum_{i=1}^{n} V_i R_{it}\text{（美元）} \tag{29}$$

其中，V_i 是投资者所持有的资产 i 在今天的固定美元价值，R_{it}是资产的（比例）历史每日收益。使用历史模拟法来计算资产组合的 Var 相当直接。这种方法假设最近的历史能代表明天所发生的事情。我们考虑种资产的收益，使用 1000 天（比方说）的历史数据来计算资产组合的价值变化，并假设 V_i 没有发生变化，然后降将这 1000 个 ΔV_{pt}值从低到高进行排列。收益分布最下端 1% 处所对应的 ΔV_{pt}值可以用来对接下来一天的 Var 进行预测（在 99% 的置信水平下）。如果在 1000 天中最低的 10 个收益为 -1000 万，-910 万，…，-220 万，那么对 1% 的尾部而言，接下来一天 Var 的预测值为 220 万。我们也可以画出 ΔV_{pt}的柱状图，得到偏度和峰度的值（这使我们能估计收益分布与正态分布之间的偏离程度）。

历史模拟方法是一种非参数方法，因为我们没有直接对方差和协方差（n 种资产）进行计算，它们隐含在历史收益和所估计的投资组合 Var 中。因此，不存在模型风险或参数估计风险。历史数据中的肥尾和自相关性（因为计算是在第 1 天、第 2 天等依次进行的）都包含在资产组合的 Var 中，由于我们已经画出了柱状图，从而不需要正态分布假设来寻找尾部截断点。借助于“Delta 逼近技巧”（将历史 R_i 作为输入变量）或更为满意地，通过历史数据（如果数据可以获得）来重新计算期权投资组合的价值，历史模拟法可以用来计算非线性头寸的 Var。也可以使用历史模拟法来计算投资组合 Var 的估计精度（例如，Butler 和 Schachter（1996）使用的是历史核密度方法）。

历史模拟法的一个缺陷在于，使用这种方法对百分位处的 Var 进行精确估计需要一个很长的时间序列数据。因为尾部事件都是“不同寻常的事件”，所以我们需要大量数据来对尾部进行稳妥估计。例如，如果我们对 1% 的下尾感兴趣，那么在我们的历史数据集合中，这种事件在 1000 个观察值中只会出现一次。另一方面，我们的历史数据集合越长，对明天 Var 的预测而言，旧数据的影响比新数据的影响就越大。Butler 和 Schachter（1996）指出基于历史模拟法的 Var 可能有很大的标准差（在 6% 和 20% 多之间）。这种方法也存在一定的弹性。例如，我们无法很容易地回答“如果某种资产的波动率比历史数据窗口所观察到的波动率增加了 10%，那么 Var 的估计值将如何变化?”然而，人们发现在某些情形中非参数方法（比如使用 1000 天的数据窗口）比参数方法（如附带 VCV 方法的 EWMA）更加精确。例如，Mohoney（1996）使用包含权益资产和外汇资产的投资组合发现，使用非参数方法所估计的投资组合 Var 与收益分布 1% 的截断点很接近，甚至 0.1% 的截断点也很接近。（反之，参数模型 EWMA 和 GARCH 在 5% 水平下的极端尾部中预测效果不是很好。）

历史模拟法的一个显而易见的主要缺陷是，结果完全依赖于所使用的特殊数据集合。数据集合可能包含未来不大可能发生的“不同寻常的事件”（例如，1992 年英镑脱离 ERM，1997—1998 年的亚洲金融危机，在美国发生的“九一一”事件）。如果是这样的话，那么用历史模拟法所估计的 Var 在预测期中是有偏的。事实上，如果我们使用的是（比方说）1000 天的历史移动数据窗口，那么这些“不同寻常的事件”的影响在 Var 预测值中会保留 1000 天。为了减轻这一问题的影响，在用历史模拟法计算投资组合的 Var 时，我们可以为更为久远的收益指定一个更小的权

重。当然，从某种意义上讲，这些问题是所有用历史数据来预测未来的方法都存在的问题，包括通常的“方差－协方差”Var方法（也包括在下面的蒙特卡罗方法）。正如我们即将看到的那样，根据原理，压力测试（stress test）明确地限制了对历史事件/数据的直接依赖性（因为“情景”是由调查者所选择的）。然而，即使是这样，历史事件也常常为调查哪些“情景”提供了指导。

“纯”历史模拟法的一个变体是自助法。正如上面所描述的那样，为了获得投资组合的Var，历史模拟法要从包含1000个每日观测值的数据集合中进行“一次一次地抽取”。因此，“抽取”过程中没有进行数据替换，抽取是按照时间顺序进行的，如第1天，第2天，等等（由此也保留了实际收益数据中的序列相关性）。在自助法中，（n种资产）T（$=1000$）个收益值按照随机的顺序进行等概率抽取。

我们可以将n种资产在某一天的收益看成一张纸上的一个维向量R。每天的历史数据分别给出了n个不同的数值，就1000天的历史数据而言，我们可以将这些纸张从1到1000进行编号。然后，我们从这1到1000的数值中随机抽取m个数（$m>n$）。例如，假设所抽取的m个随机数为$\{27,700,1,990,700,43,1,\cdots,999\}$，然后我们翻到第27，700，1，990，700，43，1，…，999页，写下这些页码上的n种资产收益R^k。我们就得到了m个向量，每个向量都表示n种资产的收益。然后针对这m天使用方程（27）来计算投资组合的价值，描绘柱状图，计算第1个百分位处的Var。因此，我们可以将种资产在1000天中的收益“复制”任意多次（在我们的情形中，总次数为m（$>n$）次）。这减缓了历史模拟法中的“数据缺乏”问题，因此从原理上讲我们能更加准确地估计投资组合的Var（尽管实践中不会总是出现这种情形，参见Butler和Schachter（1996））。

如果我们认为只有最近100天的数据才能代表将来的情形，那么取ΔV_p的某个负值作为明天第一个百分位处*Var*的预测值是非常危险的。然而，我们对这100天的数据自助次，由此得到1%尾部截断点是更具代表性的估计值。

自助法与历史模拟法有着相同的优势，主要的缺点是：初始时刻的T（$=100$）个观察值（比方说）必须能代表明天的走势。另外，数据的“更多模拟”是有代价的。因为自助法对每日收益进行随机“抽取”（从100个历史“每日”数据中进行抽取），所拟合的收益会“失去”历史收益中的序列相关性（也就是说，自助法隐含地假设收益是独立的）。然而，这个问题可以用块自助法（block－bootstrap）来加以解决：如果第x天是随机选取的，那么对这一次自助法“抽取”而言，从第$x-j$天到第$x+j$天的收益可以用来重新计算投资组合的价值。

因为监管当局允许金融机构选择自己的“内部模型”，所以历史模拟法得到了很多人的青睐，尽管这种方法对原始数据的输入要求比带参数的Var方法高，但计算量要更小一些。然而，非参数方法更有可能是带参数Var方法的一种补充而不是替代，因为后者有着更大的弹性，也更容易解释。

28.4 蒙特卡罗模拟

蒙特卡罗模拟（MCS）是一种弹性很强的参数方法，它能生成投资组合收益的完整分布，能直接“读出”合意百分位处的Var。在对包含具有非线性偿付的资产如期权的投资组合Var进行计算时，这种方法显得非常有用。它是一种参数方法，因为我们假设标的资产（如股票、债券、利率和即期外汇）的收益服从某个特殊分布，因此，在进行蒙特卡罗模拟之前，我们需要估计收益的方差和收益之间的相关系数。然而，我们不必假设收益服从多维正态分布（尽管在实践中人们经常是使用这个假设），例如我们可以用肥尾分布（如学生分布）或向左偏斜的分布（如跳跃扩散过程）来生成资产的收益。然而，需要注意的是，即使标的资产收益服从多维正态分布，一般而言，期权组合价值变化的分布也不是多维正态分布，因为标的资产收益与期权溢价之间存在非线性关系（如Black－Scholes公式）。在MCS框架内对期权进行考察，我们就能决定，在期权组合的价值变化计算过程中，是选择全局评价法（如Black－Scholes）、Delta近似法、Delta＋

Gamma 近似法，还是期权溢价的其他计算方法（如有限差分方法）。当然，当投资组合包含近价（near the money）期权时，使用线性方法或“Delta 方法”来计算 Var 可能不是很准确，因为这些资产的偿付实际上是一种非线性偿付。

压力测试（见下）和 MCS 有着相同的本质。事实上，压力测试有点像“受限”MCS，因为是分析师决定标的资产的路径，而不是使用历史数据来完全决定标的资产的路径。压力测试经常考虑标的资产的“极值”（例如，当持有货币期权时，汇率出现大幅下降），然后计算“新”极值情形中期权组合的价值变化。当然，“极端情形”的数目最好小一些。从某种意义上说，MCS 是对这种方法的一种改进，因为它能生成标的资产收益可能取值的完整序列，而这个系列值是可以用来计算期权组合的价值变化的。在下面一节中，我们先说明如何使用 MCS 来计算单一期权的 Var，然后考虑包含多种期权的投资组合的 Var。

单一资产：多头看涨期权

假设你持有以股票为标的资产的某看涨期权多头，$S_0 = 100, K = 100, r = 0.05, \sigma = 0.6$（60%），到期前还有 $T - t = 1$ 年。由 Black - Scholes 公式可知，看涨期权在初始时刻的价值（BS_0）为 25.5。该期权在 30 天的时期中 Var 为多大呢？我们可以用 MCS 来模拟 30 天中的股票价格，当然要假设股票价格服从几何布朗运动，飘移系数为 $\mu = 0.05$（参见附录 I）。

30 天之后，如果所模拟的股票价格为 S_{30}，那么由 Black - Scholes 公式可知，“新”的看涨期权溢价为 $V^{(1)} = \mathrm{BS}(S_{30}, T - 30/365, \cdots)$，我们在这里假设在 30 天的时期中其他输入量保持不变：$r = 0.05, \sigma = 0.6$。在 MCS 第一轮计算中，价值变化为 $V^{(1)} = BS_{30} - BS_0$。重复上述操作 $m(= 10\,000)$ 次就得到 $\Delta V^{(1)}(i = 1, 2, 3, ..., 10\,000)$，最后将 ΔV 按照从低到高的顺序进行排列（或画出柱状图，见图 3），左边截断点就是要计算的 Var，这里等于，也就是说，大约是初始值 $C_0 = 25.5$ 的 57%（70%）。即使股票收益服从正态分布，由于看涨期权溢价与股票价格之间存在曲率关系，所以看涨期权溢价的柱状图不是正态分布图（见图 3）。我们选择用闭型解来得出期权的价格，即 Black - Scholes 公式（也就是全值方法）。对期权溢价而言，我们则选择线性“delta 逼近法”或“delta - gamma 逼近法”来进行逼近。具体选择取决于精度和特殊评价方法的易操作性之间的权衡，我们在下一节中对此进行深入讨论。

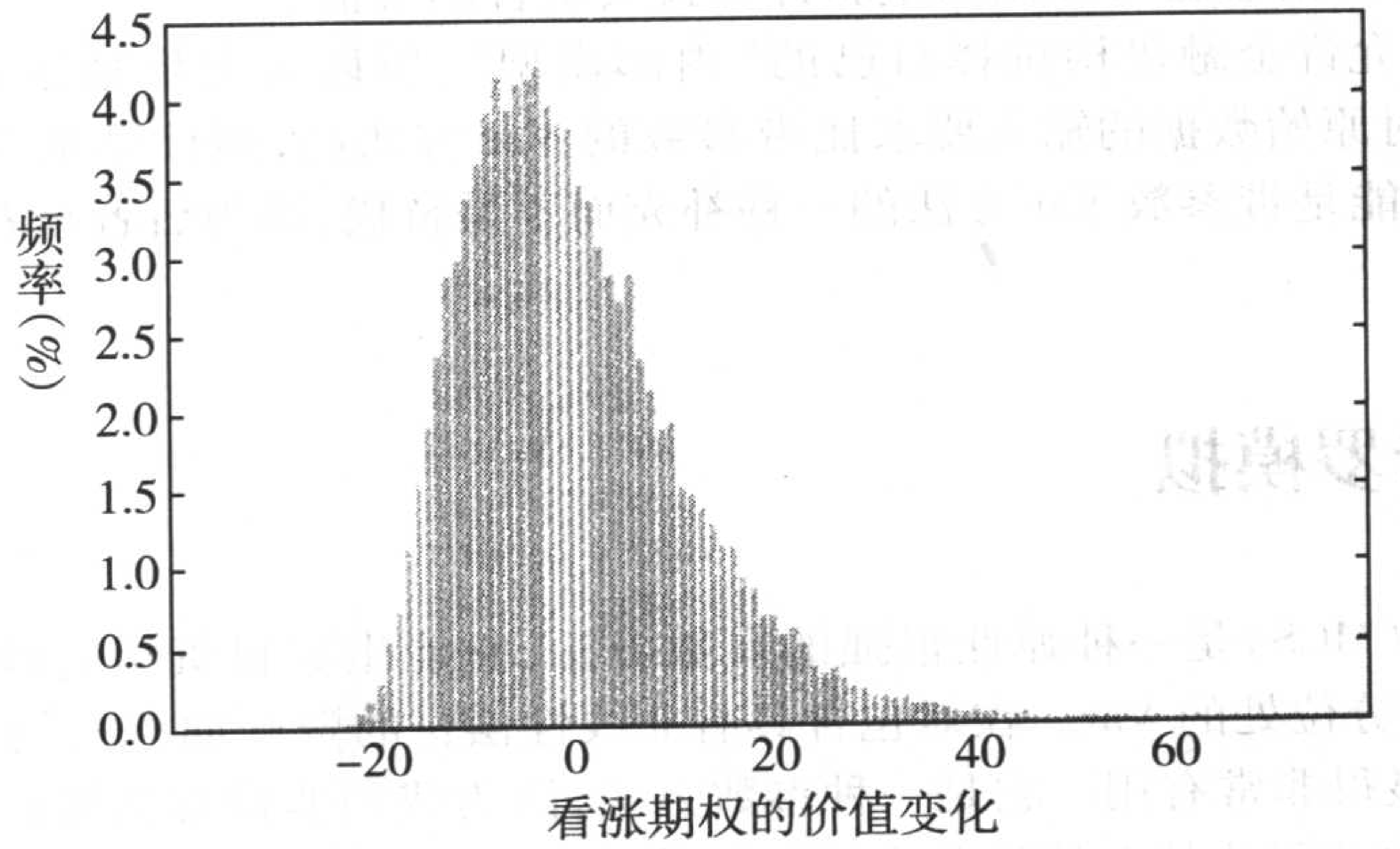


图 3　市场风险：Var

多种资产

考虑由两种资产所组成的投资组合，资产的数量为 N_i，价格为 V_i。为了使代数运算更为简便，假设我们在两种股票期权上的持有数量分别为 N_1 和 N_2，标的股票的价格分别为 P_1 和 P_2。两只股票都是国内上市的股票。如果 $N_i > 0(< 0)$，那么期权的头寸就为多头（空头）。使用 MCS

所得到的5天时期中(比方说)的Var为多大呢?蒙特卡罗方法通过如下步骤来计算第五个百分位处的Var。

资产组合的初始价值是:

$$V_0 = N_1 V_1 + N_2 V_2 \tag{30}$$

其中,V_1 和 V_2 是两种期权在 $t=0$ 时刻的初始价格。假设我们使用全局评价法和和合适的 Black - Scholes 公式来计算期权的价值。我们可以将初始期权溢价记为 $V_i = BS_0(P_i, K_i, T_i)$,$K$ 在这里是执行价格,T_i 是到期时间。我们在这里将波动率 σ 和国内利率省略了,因为我们假设这两个变量在5天之内保持不变。

使用每日波动率和相关系数的估计值,我们就可以从(多维正态)分布中进行随机抽取,生成两只股票的收益序列(为了说明问题,我们选择抽取1000次),所生成的收益序列与历史数据有着相同的相关性结构。这就是蒙特卡罗模拟。如果资产价格服从漂移项为零的几何布朗运动,那么资产1和资产2的每日(连续复利)收益为(σ_1 和 σ_2 代表每日标准差):

$$R_{1t} = \ln(P_{1t}/P_{1,t-1}) = \sigma_1 z_1 \quad z_1 \sim niid(0,1) \tag{31a}$$

$$R_{2t} = \ln(P_{2t}/P_{2,t-1}) = \sigma_2 z_2 \quad z_2 \sim niid(0,1) \tag{31a}$$

P_{it}在这里是标的股票的价格。随机变量 z_1 和 z_2 的均值为零,标准差为1,两者之间的相关系数为 ρ(也就是说,两个股票收益之间相关系数的单步(提前一天)预测值)。假设每一种资产的预期收益都为零(因为 $E(z_i)=0, i=1,2$)、收益的方差为 σ_i。假设随机变量 z_i 服从(多维)正态分布,这意味着收益服从(条件)正态分布。在 $t=5$ 天中的收益为

$$R_{i,t+5} = \ln(P_{i,t+5}/P_{it}) = \sigma_i \sqrt{t} z_i \quad i=1,2 \tag{32}$$

使用$\sqrt{T}$法则将(1000×2)收益矩阵转换为“5天前的”(1000×2)价格矩阵:

$$P_{i,t+5} = P_{i,0}\exp(\sigma_i \sqrt{5} z_i) \tag{33}$$

使用合适的 Black - Scholes 公式就可以计算出资产组合价值的10 000个“新”模拟值和同样数目的价值变化值。

$$V_{t+5}^{(1+2)} = N_1 V_{1,t+5} + N_2 V_{2,t+5} \quad 且 \quad \Delta V_{t+5}^{(1+2)} = V_{t+5}^{(1+2)} - V_0^{(1+2)} \tag{34}$$

其中,$V_{i,t+5} = \mathrm{BS}_i(P_{i,t+5}, T_i - 5/365)$。画出资产组合美元价值变化 $\Delta V_{t+5}^{(1+2)}$ 10000个模拟值的柱状图,就可以找到分布在第五个百分位处的截断点。

近似方法:Delta 和 “Delta + Gamma”

在上一节中,我们对期权使用的是全局评价法。我们也可以使用期权的 Delta(Δ_i,在 $t=0$ 时刻取值为初始股票的价格)来逼近价值变化。对看涨期权多头在30天中的 Var 而言,我们将使用全局评价法(Black - Scholes)所得到的结果与使用 Delta 近似法和 Delta - Gamma 近似法所得到的结果进行比较。对 Delta 近似法而言,

$$\Delta V_{i,t+30} = \Delta_i(P_{i,t+30} - P_i) \tag{35}$$

$P_{i,t+30}$是 $t=30$ 时股票价格的模拟值。用 Δ_i 表示赎回溢价变化的(线性)一阶近似值。包含期权 Gamma 的二阶近似值为

$$\Delta V_{i,t+30} = \Delta_i(P_{i,t+30} - P_i) + 0.5\Gamma(P_{i,t+30} - P_i)^2 \tag{36}$$

表2对这些方法进行了比较。在第1个百分位处,全局评价法给出的 Var 为17.9(下降到70%的位置),Delta 近似法给出的值为21.1(下降到83%的位置),这意味着误差很大,而 Delta - Gamma 近似法所给出的值与全局评价法所给出的值很接近,为-17.8(下降到69.9%的位置)。

在上述的蒙特卡罗模拟中,我们假设股票收益服从(条件)正态分布。但是,如果这些分布实际上具有肥尾(或者不是对称分布),那么所估计的 Var 很有可能低估真实的 Var。于是,我们需要重复 MCS,从更能代表真实数据的收益分布中获取样本。当期权偿付很“复杂”(例如,依赖于路径)且无法获得闭型解时,我们不得不在 MCS 方法框架中使用其他定价方法(参见 Cuthbertson 和 Nitzsche(2001b))。

表 2　　Var 的 delta 近似法和 delta + Gamma 近似法

		输入	
当前股票价格	100 美元	赎回溢价($t=0$)	25.5 美元
股票波动率	60% *p. a.*	执行价格 k,	100 美元
股票平均收益	5% *p. a.*	利率,r	5% *p. a.*
		Var 的计算期限	30 天
		在险价值	
百分位	**全局评价法**	**Delta 近似法**	**Delta + Gamma 近似法**
1%	17.9(70.4%)	21.1(82.7%)	17.8(69.9%)
5%	14.5(57.0%)	15.7(61.8%)	13.9(54.7%)

注:
Var 的单位为美元,括号中的数值是价值的百分比损失(即 Var/看涨期权的初始价值)。

因为 Var 计算存在上述局限性,人们也经常用压力测试来进行补充。压力测试就是估计投资组合对某个重要收益“极端运动”的敏感性。因此,它告诉你在某个特殊情形中你将损失多少,但并没有告诉你这种情形真正的发生概率。“极端运动”的选择是基于那些在经济危机时期所发生的事件(如 1987 年的股市大跌,1992 年英国货币脱离欧洲汇率机制的危机,1997—1998 年的亚洲货币危机和 2000—2003 年的股市大跌)。资产价格的方差和变动都是基于这些危机时期,相关性往往会出席急剧增加。压力测试的输入量选择取决于金融机构对投资组合主要敏感性变量可能来源的估计。显然,压力测试适用于相对简单的资产组合,因为在所选择的其他情形中,对不同的历史数据或可观察的未来事件而言,隐含相关系数的变化幅度很大。注意,压力测试的另外一个危险是,投资组合实际上可能受益于极端运动,但小幅运动则可能对其产生不利影响(如跨式期权组合多头)。

总的来说,压力测试存在一些局限性。相同的压力测试情景不可能为所有机构(和投资组合)都提供有用信息。例如,商品交易者不可能发现利率上升 300bp 的压力测试所带来的信息。人们可以转而使用情景分析法,问自己“对我的投资组合而言,在哪些情景下(对利率、汇率等而言)最坏的结果会产生?”因此,压力测试是 Var 常用算法的一个有用补充,但这些方法在实践中都需要进行大量的判断。

28.5　其他方法

简单地说,Var 试图在给定的概率下用一个数值来反映资产组合的市场风险。这个“数值”可以用多种方法来进行计算,数值的大小取决于所选择的建模过程。例如,在所谓的参数方法中,人们经常假设收益服从正态分布,经常使用波动率和相关系数的特殊预测值,也可以考虑分散化效应。当资产组合包含期权时,人们经产使用蒙特卡罗模拟这种参数方法。显然,这些方法都有一定的“模型误差”,因为所估计的方差和协方差可能是错误的。与之相对照的是,在非参数“历史模拟法”中,给定真实的历史收益数据后,我们就能“观察到”资产组合价值本应该具有的变化。这种方法也经常给出单一的 Var,尽管它很容易给出资产组合收益的完整分布。这种方法需要大量的数据,结果有时高度依赖于所选择的历史数据。

我们来考察上述方法所存在的问题,从某种意义上说,这些方法是近期数据的一种“平均”。例如,对市场价值为 $V=6.06$ 亿美元且 $\sigma=1\%$(每日)的资产组合而言,第 5 百分位处的每日 Var 为 1000 万美元($=V_0 1.65\sigma$)。这意味着 20 天中有 19 天的预期损失不会超过 1000 美元,预期损失超过 1000 万美元的至多 1 天。然而,对这一天的真实损失而言,Var 没有告诉我们答案。如果

收益服从（条件）正态分布，那么0.5%的左尾部截断点为3.2σ，即Var为1940万美元。因此，我们可以“很自信”地说损失不会超过2000万美元。但是，尽管如此，我们还是无法确切地知道真实损失为多大，原因在于：

* 即使在正态分布假设下，大额损失发生的概率也很小。

* 真实收益比正态分布有着更肥的尾部，所以更大更频繁的损失（平均而言）将会发生（相对于正态分布假设下的计算结果而言）。

* 相关系数和波动率的变化非常大。例如，在1993年的第一季度中，日经225指数每日收益和英国金融时报100指数每日收益之间的相关系数从+0.9变化到-0.9（Jackson、Maude和Perraudin（1997）），日经225指数在1995年第一季度的每日波动率在0.7%到1.8%之间变动。

更坏情形的情景分析（WCS）

解决（至少可以减缓）上述“损失大小问题”的一种方法是对更坏情形进行情景分析（WCS）。我们在这里不是简单地计算第5个百分位截断点处的Var，而是只考察下尾中的结果，并用这些结果来计算预期的最坏结果（Boudoukh、Richardson和Whitelaw（1995））。因此，这种方法试图回答如下问题：“假设我们处于尾部情形中，那么接下来H个交易日最大损失的平均值是多少？”尽管Var关注的是某个特殊损失发生的频率（例如，1%水平下1百万美元的Var意味着，在100天中只有1天的损失会超过1百万美元）。WCS的基础是蒙特卡罗模拟，这种方法所给出“风险值”比Var大，因为它是根据下尾极值收益的分布来进行计算的。

例如，Boudoukh、Richardson和Whitelaw（1995）所报告的1% Var为-2.33%（标准差），但*WCS*平均损失为-2.51%，对“更坏情形”的结果而言，1%尾部值为-3.72%。使用WCS分析的一个主要问题是，相对于蒙特卡罗模拟法而言计算成本有所增加，这种方法还假设（在由多种资产所组合的投资组合中）使用历史数据所估计的相关系数适用于收益的极端运动，正如分布的尾部所表示的那样。

极值理论

关注分布尾部的另外一种方法是极值（EV）方法。这种方法在自然科学中广为使用，人们用它估计某一置信度下的最大损失（例如大坝和原子能核电站出现问题）。从宽泛意义上讲，EV方法是历史模拟法与分布尾部形状估计的结合。与历史模拟法不同的是，极值法只使用下尾的数据来估计分布尾部的形状，不会施加与分布有关的其他假设。Danielsson和deVries（1997）认为，在小于5%的百分位处，VCV-Var方法倾向于低估尾部事件发生的可能性，历史模拟方法有高估的倾向，而EV方法则表现得很好。因此，对人们所感兴趣的尾部概率而言，EV方法似乎要更为准确一些。

前面讨论的所有方法（如方差-协方差参数方法、历史模拟方法、压力测试、蒙特卡罗模拟和*EV*理论）一般都假设资产组合的持有数量在所考虑的时期中保持不变。如果Var的计算时期更长一些，那么这一限制条件就会变得更加不现实。例如，投资者可能无法在一天之内清算所有的头寸（特别是在经济危机期间），但对比10天更长的时期而言，上述假设不再成立。所有方法都遗漏掉的一个重要问题是“反弹时期”问题。在市场定价的基础上，10天累计损失的估计值可能很大，但是，如果市场在20天内就完全恢复过来了，那么30天内的实际累计损失接近于零。因此，一种很有价值的做法是，不使用$\sqrt{T}$规则来考察Var，而是对更长时期中的波动率直接进行估计（例如，收益可能呈现出均值反转特征）。然而，这需要大量数据来估计（比如说）月度收益的波动率。

总之，Var计算过程中所使用的统计方法和实际时间计算方面都会给人留下深刻的影响。然而，到目前为止一个很明显的事实是，在对市场风险（也就是说，将信用风险、操作风险等排除在外）的度量而言，任何方法都需要进行大量的判断。

28.6 小结

* 如果假设收益服从（多维）正态分布且资产组合的价值与资产收益线性相关，那么 Var（在任意给定的百分位处）很容易用 Delta－Normal 方法或方差－协方差（VCV）方法计算出来。

* 如果收益独立（同）分布，那么我们可以用$\sqrt{T}$规则来计算不同时期中的波动率（即 $\sigma_T = \sqrt{T}\sigma$，$\sigma$ 在这里是指每日波动率的预测值）。

* 映射缓和了“大量数据需求及数据度量”的问题，使得 VCV 方法可以用于一些非线性工具（如债券）。单一指数模型（SIM）可以用来简化某个国家权益资产的 Var 计算。VCV 方法对股票、债券、外汇、互换和期货都非常有用。

* 包含“大众型”欧式期权的投资组合的 Var 可以用用蒙特卡罗模拟（MCS）来进行准确计算，人们也常用压力测试来对此进行补充。

* 历史模拟法对下尾截断点（例如，第一百分位和第五百分位之间的截断点）的预测准确度非常高，但这种方法需要较长的时间序列数据，且分析没有弹性（例如，很难进行敏感性分析），这与标准的方差－协方差参数方法形成对比。

* 与 WCS 分析和 EV 方法类似，对分布尾部进行考察不可能产生“很大的价值”，因为方法的复杂性也所有增加。

* 一些形式的压力测试为根据参数和非参数 Var 方法所得到的市场风险提供了一种补充看法。但在确定“真实”的系统风险时，所有方法都需要进行大量判断。

附录 I：蒙特卡罗分析与 Var

看涨和看跌期权溢价都是标的资产价格的非线性函数。蒙特卡罗模拟可以用来计算那些包含期权（和其他资产）的投资组合的 Var。如果我们希望计算包含两种或多种资产的投资组合的美元价值，那么我们就必须考虑资产收益之间的相关性。假设标的资产的（连续复利）收益服从多维正态分布，因此资产价格服从多维对数正态分布。我们用两种资产的情形来说明这种方法，但很容易将其扩展为 n 种资产的情形，也可以选择合适的统计软件来进行编程。

假设资产收益 R_i 服从多维正态分布，也就是说，$R \sim N(0, \Sigma)$，假设我们已经有收益方差－协方差矩阵的估计值/预测值，由此可以计算相关系数矩阵 C。对（2×2）的情形而言，我们有

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \tag{A1}$$

且

$$C = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \tag{A2}$$

其中，$\sigma_{12} = \sigma_{21}$，相关系数为 $\rho = \sigma_{12}/\sigma_1\sigma_2$。简单的矩阵运算就可以使我们在 C 矩阵和（2×2）矩阵“A”之间建立映射关系，从而

$$C = AA' \tag{A3}$$

“A”是下三角矩阵：

$$A = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix} \tag{A3}$$

矩阵 A 是矩阵的 Choleski 分解。很容易验证 $C = AA'$。进行 Choleski 分解的原因在于，如果 ε（2×1）由两个独立的 $N(0,1)$ 变量组成（即均值为零，方差为 1，协方差为零，$\varepsilon \sim N(0,I)$），那么我

们可以生成一组服从标准正态分布的向量 Z：

$$Z = A\varepsilon \tag{A5}$$

Z 中变量的均值为零，方差为 1，相关系数为 ρ。因此，Z 与相关系数矩阵 C 有着相同的相关系数结构。一旦生成了 z_i，我们就可以用下式来导出 t 期中两种资产的收益和价格：

$$P_t^{(i)} = P_0^{(i)}\exp(\sigma_i\sqrt{t}z_i) \tag{A6a}$$

$$R_{i,t} = \ln(P_t^{(i)}/P_0^{(i)}) \tag{A6b}$$

现在，我们很容易证明在从 $t=0$ 到 t 的时期中，资产收益（方程）的标准差为 $\sigma_i\sqrt{t}$，相关系数为 ρ。为了看清这一点，注意，由方程（A4）和（A5）就可以得到

$$z_1 = \varepsilon_1 \tag{A7a}$$

$$z_2 = \rho\varepsilon_1 + \sqrt{1-\rho^2}\varepsilon_2 \tag{A7b}$$

变量 z_1 和 z_2 有着共同的一项 ε_1。这是 z_1 和 z_2 之间相关性的产生根源，因此也是两种资产收益之间相关性的产生根源。由方程（A6a）和（A6b）可知

$$R_{1t} = \sigma_1\sqrt{T}z_1 \tag{A8a}$$

$$R_{2t} = \sigma_2\sqrt{T}z_2 \tag{A8b}$$

为了看清模拟收益的均值为零且相关系数为 ρ 这一事实，注意 $Ez_1 = Ez_2 = 0$，因为 $\sigma E(\varepsilon_1) = E(\varepsilon_2) = 0$。于是，

$$\mathrm{Var}(z_1) = \mathrm{Var}(\varepsilon_1) = 1 \tag{A9a}$$

$$\mathrm{Var}(z_2) = \rho^2\mathrm{Var}(\varepsilon_2) + (1-\rho^2)\mathrm{Var}(\varepsilon_2) = 1 \tag{A9b}$$

$$\mathrm{Cov}(z_1,z_2) = E(z_1z_2) = \rho E(\varepsilon_1^2) + \sqrt{1-\rho^2}E(\varepsilon_1\varepsilon_2) = \rho \tag{A9c}$$

因为 $\mathrm{Cov}(z_1,z_2)=0$，所以 T 天时期中收益的方差为

$$\mathrm{Var}(R_i) = \sigma_i^2T\mathrm{Var}(z_i) = \sigma_i^2T$$

使用方程（A8a）和（A8b）就可以证明收益（在 T 天内）之间的相关系数与每日收益的相关系数相同：

$$\rho_T = \mathrm{Cov}(R_{1,t},R_{2,t})/\sigma_1\sqrt{T}\sigma_2\sqrt{T} = \rho \tag{A9d}$$

最后一个细微差别值得我们注意。我们在上面使用 MCS 来生成单期收益，然后使用$\sqrt{T}$规则来计算 $t+5$ 时刻的价格。此外，使用下面的任何一个递归式子都可以计算出 P_{t+5}（这是对几何布朗运动的离散近似）：

$$P_t = [1+\mu\Delta t+\sigma\varepsilon_t\sqrt{\Delta t}]P_{t-1} \tag{A10a}$$

$$P_t = P_{t-1}\exp[(\mu-\sigma^2/2)\Delta t+\sigma\varepsilon_t\sqrt{\Delta t}] \tag{A10b}$$

我们通常取 σ 为年度标准差，μ 为年度平均收益，Δt 是时间步长（以年为单位（例如，1 天 = 1/365 年））。在上述简化后的例子中，我们在第二个方程中取 σ = 每日波动率，Δt = 1 天，$\mu=0$。第一个方程采用“乘法形式”，包含 $N(0,1)$ 随机项，它只是 P 对数正态分布的一种近似，而第二个方程是任意时间步长 Δt 中对数正态分布的精确表示。

事实上，如果我们只是对特殊时刻 T 的 P 值感兴趣，那么“一步”就可以计算出 P_T：

$$P_T = P_0\exp[(\mu-\sigma^2/2)T+\sigma\varepsilon_t\sqrt{T}] \tag{A11}$$

实践问题

绝大多数统计软件都能产生服从均匀分布的伪随机数，也就是说，以相同的概率在 0 和 1 之间进行取值。可以使用多种方法将这些数转变成服从标准正态分布的随机数。例如，如果是服从均匀分布的随机数，那么 $Z_i = \sum_{i=1}^{12}\varepsilon_i - 6$ 就是服从正态分布的随机变量。然而，这种方法产生了太多与均值接近的数值，但通过其他“变换方法”（如 Box – Muller 变换）可以减缓这个问题的影响。

在 MCS 中存在两个关键的“数字”问题：所模拟的“伪随机数”实际上不是随机数；无法对这

些数值进行 Choleski 分解。随机数生成器实际上是一种定性算法，以某个特殊的“种子数字”为起点所生成的似乎是随机数（能通过独立性检验等）。对给定的“种子数字”，随机数集合可以重复生成。问题在于，随着数字生成“数目”的增加，随机数生成器可能选择一个已经使用过的种子。这就导致了重复生成或随机数“循环”，蒙特卡罗模拟的准确性似乎会增加。另一个问题是，标准 MCS 倾向于产生“聚类”数字，所以增加观察次数无法提供新信息，当“模拟次数”很小时，这会使结果出现偏差。然而，所谓的准蒙特卡罗模拟（QMC）方法或低差序列可以用来生成随机数，同时又可以避免“聚类”现象的产生。因此，MCS 变得越来越有效，较小的“模拟次数”就能产生相当准确的结果。这种方法的思想在于，“准随机”样本“记住了”以前的样本并试图使自己远离以前的样本，从而能“填充”样本空间却又可以避免聚类现象。有很多方法可以生成准随机样本，读者可以参考 Clewlow 和 Strickland（1998）。

第二个问题在于，为了进行 Choleski 分解，所估计的方差 - 协方差矩阵 Σ 必须是半正定矩阵。这就要求观察值的数目大于协方差矩阵的维数且收益之间不存在完全共线性问题。如果 Σ 很大，那么近似完全共线性是可能的，这就是为什么我们要通过“映射”来减小收益取值的原因。

附录 II：单一指数模型

我们希望说明如何用单一指数模型（SIM）来计算投资组合的方差：

$$\sigma_p = \left[\sum_{i=1}^{n} \omega_i\beta_i\right]\sigma_m = \beta_p\sigma_m \tag{A1}$$

其中，σ_p 是投资组合的标准差，σ_m 是市场组合的标准差，$\beta_p = \sum {}_i\omega_i\beta_i$ 是投资组合的“beta”。SIM 可以用方程（A2）和假设（A3）来表示：

$$R_i = \alpha_i + \beta_i R_m + \varepsilon_i \tag{A2}$$

$$E(\varepsilon_i) = 0 \tag{A3a}$$

$$E(\varepsilon_i\varepsilon_j) = \sigma^2_{\varepsilon_i}(i = j)\text{，否则为零} \tag{A3b}$$

$$\mathrm{Cov}(R_m, \varepsilon_t) = 0 \tag{A3c}$$

于是，

$$ER_i = \alpha_i + \beta_i ER_m \tag{A4a}$$

$$\sigma_i^2 = \beta_i^2\sigma_m^2 + \sigma^2_{\varepsilon_i} \tag{A4b}$$

$$\sigma_{ij} = \beta_i\beta_j\sigma_m^2(i \neq j) \tag{A4c}$$

投资组合的预期收益和标准差（方差）定义如下：

$$ER_p = \sum_i \omega_i ER_i \tag{A5a}$$

$$\sigma_p^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \omega_i\omega_j\sigma_{ij} \quad (i \neq j) \tag{A5b}$$

投资组合方差的上述公式需要 n 个方差和 $n(n-1)/2$ 个协方差。当 $n = 150$ 时，“输入”数量达到 11 325 个。为了计算 ER_i，我们还需要 $n = 150$ 个 α_i 值，从而“输入”数量达到 11 475。为了减少输入数量，我们可以使用 SIM。将（A4a）代入（A5a）就得到

$$ER_p = \sum_i (\omega_i\alpha_i + \omega_i\beta_i ER_m) = \alpha_p + \beta_p ER_m \tag{A6}$$

其中，$\alpha_p = \sum_i \omega_i\alpha_i$，$\beta_p = \sum_i \omega_i\beta_i$。将（A4）中的 σ_i^2 和 σ_{ij} 代入（A5b）就得到

$$\begin{aligned}\sigma_p^2 &= \left[\sum_{i,j}^{n} \omega_i\omega_j\beta_i\beta_j\right]\sigma_m^2 + \sum_{i=1}^{n} \omega_i^2\sigma^2_{\varepsilon_i} \\ &= \left(\sum \omega_i\beta_i\right)\left(\sum \omega_j\beta_j\right)\sigma_m^2 + \sum \omega_i^2\sigma^2_{\varepsilon_i}\end{aligned} \tag{A7}$$

$$\sigma_p^2 = \beta_p^2\sigma_m^2 + \sum \omega_i^2\sigma_{\varepsilon_i}^2 \tag{A8}$$

方程（A8）可以解释为

$$\textbf{投资组合的总风险} = \textbf{市场风险} + \textbf{特殊风险} \tag{A9}$$

在 SIM 中，ER_p 和 σ_p^2 的计算需要 α_i、β_i 和 σ_{ε_i} 的 $3n$ 个估计值，还需要 σ_m^2 和 ER_m 的估计值，从而“输入”总数为 $3n+2$（当 $n=150$ 时，“输入”数量为 452 个）。这是对（A5a）和（A5b）的极大简化。特殊风险仅仅是回归（A2）的（加权）残差平方和。如果我们没有持有分散化投资组合，那么（A8）的最后一项不等于为零，我们必须将其纳入到 σ_p^2 的计算之中。然而，我们在这里假设对投资组合的所有股票而言，有 $\omega_i = 1/n$，于是

$$\sigma_p^2 = \beta_p^2\sigma_m^2 + n^{-1}\left(\sum \sigma_{\varepsilon_i}^2/n\right) \tag{A10}$$

括号中的最后一项是平均而言的特殊风险 $\bar{\sigma}_{\varepsilon_i}$。因此，最后一项为 $(1/n)\sigma_{\varepsilon_i}$，当 $n\to\infty$ 时，这一项趋于零（随机选择 $n(=25)$ 只股票就可以保证这一项足够的小）。因此，在分散化良好的投资组合中，

$$\sigma_p = \left[\sum_{i,j}^{n} \omega_i\beta_i\right]\sigma_m = \beta_p\sigma_m \tag{A11}$$

如果 V_i 是资产 i 的持有头寸，那么投资组合的价值为 $V_p = \sum_{i=1}^{n} V_i$，$\omega_i = V_i/V_p$。投资组合的 Var 可以用两种方式来表示：

$$\mathrm{Var}_p = V_p(1.65)\sigma_p = \sum_i V_i(1.65\beta_i)\sigma_m = \sum_{i=1}^{n} \mathrm{Var}_i \tag{A12}$$

其中，$\mathrm{Var}_i \equiv V_i\beta_i(1.65)\sigma_m$。也可以等价地表示为

$$\mathrm{Var}_p = \sum_{i=1}^{n} \mathrm{Var}_i = (ZCZ')^{1/2} \tag{A13}$$

其中，$Z = [\mathrm{Var}_1, \mathrm{Var}_2, \cdots, \mathrm{Var}_n]$，$C$ 是单位矩阵。RiskMetrics™给出了 σ_m 的估计值，用户需要知道 V_i 的值和 β 的估计值。SIM 中一个关键假设是 $E(\varepsilon_i\varepsilon_j)=0$，也就是说，$i$ 公司（或行业）冲击的方差与 j 公司冲击的方差之间不会同期相关。然而，正如人们在每日收益中所发现的那样，如果 ε_i 和 ε_j 之间的相关系数为正但很小，那么随着 n 的增加，特殊风险也会出现急速下降。

一个附带的值得我们注意的问题是在 SIM 的假设下，股票之间的相关系数可以根据 β 的估计值（和方差）来进行计算：

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i\sigma_j} = \frac{\beta_i\beta_j\sigma_m^2}{\sigma_i\sigma_j} \tag{A14}$$

这里的 σ_i 和 σ_j 由（A4b）给出。注意，如果 SIM 的拟合效果很好，从而 $\sigma_{\varepsilon i}=0$，那么 $\rho=1$，因为每一种资产的收益都由 R_m 来决定。然而，在 RiskMetrics™中，相关系数 ρ_{ij} 是直接进行计算的，因为 RiskMetrics™的目的是为股票、债券、外汇等的相关系数预测提供一种普遍适用的方法。注意，所谓的市场模型放松了 SIM 的如下假设：$\mathrm{Cov}(\varepsilon_i,\varepsilon_j) = E(\varepsilon_i\varepsilon_j) = 0$。但是，如果这样做的话，那么我们的投资组合风险简化表达式（A11）就不成立。一旦 $E(\varepsilon_i\varepsilon_j)\neq 0$，$\sigma_p$ 的计算就需要所有的协方差。

第29章

波动率与市场微观结构

目的

- 计算资产收益的波动率，考察影响波动率的各个经济变量。
- 明确时变风险溢价对预期资产收益的重要性，时变风险溢价可以用 ARCH 和 GARCH 过程来进行建模，这里的资产包括股票、债券和外汇等。
- 考察外汇市场的结构，它会对买卖差价和即期汇率产生影响。
- 说明如何使用调查数据来对理性预期假说和不同类型交易者对价格变化的影响进行评估。
- 借助高频交易数据通过技术交易法则来获取“超额利润”。

正如我们在前面章节中所注意到的那样，ARCH 和 GARCH 过程可以用来对那些取决于时变方差和协方差的资产收益进行建模（如标准 CAPM 模型和 SDF 模型）。在这一章里，我们对那些适用于股票、债券和外汇收益的方法进行总结。我们将说明风险溢价中的持续性如何导致股票价格的大幅波动（正如在真实数据中所观察到的那样）。风险溢价持续性的估计对股票收益预测方程中所包含的其他经济变量非常敏感，如红利价格比、无风险利率和市场交易量。因此，我们将考察预期收益、时变方差和协方差之间关系的稳健性。我们将用例子来说明，多变量 ARCH - GARCH 模型在用股指期货进行动态套期保值方面是如何应用的，如何用它们来对不同期限债券的时变风险溢价进行度量，如何用它们来对股票和债券收益之间的协方差进行计算。

本章余下部分致力于市场微观结构的讨论。从传统来看，这一领域涉及短期内价格和收益不同类型的交易结构所产生的影响。波动率显然是故事的一个部分，但这一领域也考察买卖差价的决定因素和“交易性变量”（如指令流，交易规模）是否影响价格。我们也在这个主题下对调查数据进行讨论，投资者如何预测未来价格与理性预期和危机消除之间存在直接联系。最后，技术交易法则（如图表主义、滤波器法则和遗传算法）被经常应用于高频数据，因此，我们将考察在进行多次“赌博”后这些法则还是否能获得“异常利润”的问题。

29.1 波动率

我们来回顾一下与收益波动率有关的一些“典型事实”，这里的收益计算时期至多不超过一个月（如一日内的收益、每日收益、每周收益和每月收益）。股票和即期汇率似乎经历了一些变化不等的时期，有些时期中的变化较大，有些时期中的变化较小。此外，当收益变化很小（大）时，收益倾向于在一段时间内都变化很小（大）。因此，股票收益的波动率在短期内（如一个月）具有如下性质：

（ⅰ）波动率具有时变性（即平静时期和动荡时期都存在）。

（ⅱ）波动率存在聚集效应：在收益出现大（小）的变化之后倾向于出现大（小）的收益变化（波动率具有自回归性质和持续性）。

（ⅲ）收益的无条件分布具有“肥尾”。

（ⅳ）各种资产收益之间的（条件）相关系数似乎会随着时间而发生变化。

我们需要一些模型来解释这些“典型事实”，因为时变方差和协方差在资产组合配置和风险度量（如 Var）方面非常重要。

波动率不是直接进行度量的。基于市场视角的波动率度量可以从 Black - Scholes 期权定价方程（或其他闭型解）“导出”。我们所能观察到的期权价格为 $P_t = f(S_t, \sigma_t, r_t, T-t, K)$，$f$ 在这里是一个非线性函数。除 σ_t 之外的所有变量在 t 时刻都是已知的，因此我们可以在期权的有效期内“逆向”推导波动率的市场预测值（这就是众所周知的隐含波动率）。如果期权（如股票期权、股票指数期权、期货期权等）在流动市场上进行交易且期权定价具有闭型解，那么我们就可以求出标的资产的隐含波动率序列。一般而言，隐含波动率可以作为一个基准来与波动率的统计指标进行比较，这是本章的主题。

波动率的一个度量指标，范围估计值（range estimator），定义如下：

$$h_t = \ln(high_t / low_t) \tag{1}$$

这里的“high”和“low”是在所考虑的时期中（如一个交易日内）观察到的价格。如果波动率存在自相关性，那么我们就可以对如下方程进行估计：

$$h_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j h_{t-j} + \varepsilon_t \tag{2}$$

（每日）收益的平方 $h_t = (R_t - \mu)^2$ 也可以用来估计波动率，就月度或更低的频率而言，一种合理的近似方法是假设 $\mu = 0$。显然，方程（1）中只使用一天内的两个数据点。人们也可以选择用每小时收益平方的样本平均值来度量（每日）波动率。

股票收益的另外一个简单而又很常用的度量指标是使用指数加权移动平均 EWMA。波动率在这里具有时变性，使用历史数据的移动窗口来对其进行预测。例如，在对每日波动率进行预测时，EWMA 模型在 Var 的计算中被广泛使用：

$$h_{t+1} = \lambda h_t + (1-\lambda) R_t^2 \tag{3}$$

可以对系数 λ 进行估计或选择一个 λ 值来使样本外预测误差达到最小（例如在 RiskMetrics™ 中，我们为每日数据选择的是 $\lambda = 0.94$）。通过后向递归，方程（3）就意味着波动率的预测值是收益平方历史值的长期加权平均，因此波动率具有持续性。

ARCH - GARCH

最简单的 GARCH 模型是：

$$R_{t+1} = \mu + \varepsilon_{t+1} \tag{4a}$$

$$h_{t+1} = \alpha_0 + \alpha_1 h_t + \alpha_2 \varepsilon_t^2 \tag{4b}$$

$$h_{t+1} \equiv E_t(R_{t+1} - E_t R_{t+1})^2 = E(\varepsilon_{t+1}^2 \mid \Omega_t) \tag{4c}$$

h_{t+1} 是收益 R_{t+1} 的条件方差，等于 $E(\varepsilon_{t+1}^2 \mid \Omega_t)$，这里的 Ω_t 是 t 时刻的信息。模型（4b）假设条件方差具有自回归性质，这就产生了波动聚集（volatility cluster）效应。无条件方差为常数且等于：

$$\mathrm{Var}(\varepsilon_{t+1}) = \alpha_0 / [1 - (\alpha_1 + \alpha_2)] \quad ((\alpha_1 + \alpha_2) < 1) \tag{5}$$

当 $(\alpha_1 + \alpha_2) = 1$ 时，无条件方差没有定义，这个模型就是众所周知的 IGARCH(1,1)。注意，如果 $\mu = 0$，那么 EWMA 模型就是 IGARCH 模型。（4a）中误差项的等价表示是 $\varepsilon_{t+1} = \sqrt{h_{t+1}} u_{t+1}$，这里的 u_{t+1} 独立同分布，均值为零，方差为 1。“标准化后的”残差 $\hat{v}_t = \hat{\varepsilon}_{t+1} / \sqrt{\hat{h}_{t+1}}$ 是一个随机变量，可以对其正态性进行检验（如 Bera - Jarque 检验）。实践中，对股票收益而言，人们经常发现 $\hat{v}_t$ 具有尖峰（即厚尾）特征，但程度要比 ε_t 小。使用极大似然估计法来对 GARCH(p,q) 模型进行估计

已经是一种很标准的方法，许多计量经济学软件包（如 TSP，EVIEWS，PCGIVE，MFIT，RATS，GAUSS - FanPac）都含有相关程序。

29.2　波动率的影响因素

Schwert（1989）考虑了时变波动率的可能来源。他考察了股票收益的条件波动率在多大程度上取决于股票收益的历史波动率和其他经济变量（基本经济变量）的波动率，如债券波动率和实际产出的波动率。如果所感知的风险具有持续性，那么今天的风险增加会导致未来多个时期所感知风险的增加。由理性定价公式 RVF 可知，股票价格取决于风险溢价。因此，如果风险具有持续性，那么所感知风险的微小增加都会导致当前股票价格出现大幅下跌。

Schwert（1989）对股票收益的条件波动率进行了考察，也就是说，股票收益的波动率以所得到的股票收益最佳预测值为条件。如果我们将股票收益的最佳预测值记为 $E(R_{t+1}|\Omega_t)$，那么 $\varepsilon_{t+1}=(R_{t+1}-E_tR_{t+1})$ 就是条件预测误差。因为 $E_t\varepsilon_{t+1}=0$，所以收益预测误差的条件方差为 $\mathrm{Var}(\varepsilon_{t+1}|\Omega_t)=E_t(R_{t+1}-E_tR_{t+1})^2$。Schwert 用一个简单指标来度量条件波动率。他假设股票月度收益的最佳预测值 R_{t+1} 可以由 AR 模型来给出（我们将月度虚拟变量排除在外）：

$$R_{t+1}=\alpha_0+\sum_{j=0}^{m}\alpha_jR_{t-j}+\varepsilon_{t+1} \tag{6}$$

Schwert 发现方程（6）的残差 $\hat{\varepsilon}_{t+1}$（的绝对值）具有序列相关性，并用 AR 过程来对其进行建模：

$$|\hat{\varepsilon}_{t+1}|=\sum_{j=1}^{s}\rho_j|\hat{\varepsilon}_{t-j+1}|+u_{t+1} \tag{7}$$

这是预测误差一种（简单）形式的自回归条件异方差模型（ARCH 模型）。我们由 ARCH 回归就可以得到 ρ_j 的估计值。由方程（7）所得到的预测值 $\varepsilon_{t+1}^{*}=\sum_{j=1}^{s}\hat{\rho}_j|\hat{\varepsilon}_{t-j+1}|$ 给出了股票收益条件标准差的时变估计值。Schwert 也用上述方法来计算其他经济序列的条件波动率，将方程（6）和（7）用于股票收益（市场指数的收益）、债券收益、通货膨胀、短期利率、工业产出增长率和货币增长率的月度数据。

方程（7）中 ρ_j 的和是波动持续性的一个度量指标。例如，当 $\rho=0.95$ 时，t 时刻的单位冲击（即 $u_t=1$）在 12 个月后对 ε_{t+12} 的影响为 0.54，在 24 个月后对 ε_{t+24} 的影响为 0.3。然而，如果 $\rho=0.5$，那么 u_t 对 ε_{t+12} 的影响几乎为零。因此，较大的 ρ 值意味着一个“长时间记忆”。Schwert 对 1859—1987 年期间美国的经济序列进行了考察，发现了波动持续性现象，ρ 的和介于 0.8 和 0.85 之间。因此，股票收益的波动持续性可以用基本经济变量的波动持续性来反映。

现在考虑股票收益条件波动率 $\hat{\varepsilon}_{t+1}$ 与基本经济变量 ε_{it} 条件波动率之间的可能关系，这里的 i 可以是产量、债券利率和通货膨胀率等。Schwert 将股票收益波动率 $|\hat{\varepsilon}_t|$ 对其滞后值和基本经济变量 $|\varepsilon_{it}|$ 的滞后值进行了回归。任何“逆向影响”（也就是说股票波动率对产出等基本经济变量波动率的影响）都能在以产出波动率为因变量的方程中得以体现。事实上，Schwert 大致估计了股票收益波动率方程和向量自回归 Var 系统中的“逆向回归”效应。

Schwert 的结论不是很一致。他几乎没有找到基本经济变量（如产量、通货膨胀等）的波动率对股票收益波动率有明显影响（影响随着时间并不稳定）的证据。然而，利率和公司债券利率的波动率对股票波动率的影响在统计上是显著的。另外，其他“没有波动”的经济变量也确实对月度股票收益的条件波动率有影响。这些变量包括债务权益比率（杠杆率），就月度交易活动而言，这种影响是一种正面影响。交易活动可以用交易量或买/卖指令数量的增长率和月内交易天数来进行度量。股票的波动性在衰退时期比经济繁荣时期更强。“逆向回归”结果表明，有微弱的证据能证明股票收益波动率对产出波动率有更强的解释力。在 Schwert 的研究中，股票收益波动率的大部分运动无法用所考察的经济变量来进行解释，所报告的回归 R^2 经常落在 0 ~ 0.3 的区域内。因此，股票收益的许多月度波动都可以归因于“消息”。

收益和波动率

如果股票综合指数的风险溢价可以很好地用条件方差的 ARCH 模型来进行度量，那么未来风险溢价是部分可预测的。方差持续性意味着，今天的方差增加将导致未来各期股票认知风险的增加，由 RVF 可知，这也会导致股价的大幅变化。Poterba 和 Summer(PS)(1988)是最早对“投资者的认知风险变化是否足够的大，以至于能解释实际观察到的大幅股价变化”进行考察的文章之一。如果他们的模型能解释股票价格的实际运动，那么基于理性行为的股票价格就不会出现过度波动现象。

PS 对 RVF 进行了线性逼近，他们能准确地说明股票价格如何对股票认知风险的突然增加产生反应。这种反应在很大程度上取决于条件波动率的持续性程度。如果波动持续性为 0.5，那么 PS 的计算结果是：波动率 1% 的增加会导致股票价格 1.2% 的下降。然而，如果持续性为 0.99，那么波动率每增加 1%，股票价格就会下降 38%，这种效应相当强烈。假设股票价格在很短的时间段经常发生急剧变化，那么为了使 PS 模型“拟合事实”，我们需要找到一个很强的波动持续性。

PS 将 RVF 中的所有变量都处理成实际变量，P_t 是 1928—1984 年期间实际的 S&P 指数（相对名义指数而言）。因为 *PS* 希望关注风险溢价对 P_t 的影响，所以他们假设实际红利增长率为常数 g，实际无风险利率为常数。跨期 CAPM 模型（Merton(1973，1980)）认为，市场投资组合的风险溢价与权益收益预测误差的条件方差 $h_{t+1} \equiv E_t\sigma^2_{t+1}$ 成比例，因此

$$E_tR_{t+1} = r_t + \lambda h_{t+1} = r_t + rp_t \tag{8}$$

其中，λ 是风险的市场价格，$rp_t = \lambda h_{t+1}$。在 Merton 的跨期 CAPM 模型中，λ 取决于不同消费者相对风险规避系数的加权平均（我们在这里假设它为常数）。PS 假设条件波动率可以用 ARMA 模型来表示，这里假设它服从过程：

$$h_{t+1} = \alpha_0 + \alpha_1 h_t + v_{t+1} \quad (0 > \alpha_1 > 1) \quad (0 > \alpha_1 > 1) \tag{9}$$

v_{t+1} 是白噪音误差。通过 v_{t+1} 项所提供机制，将其从正变为负，我们就可以（随机地）从一个高波动时期转换到一个低波动时期。大的 α_1 值意味着波动持续性很强。如果 h_t 服从 $AR(1)$ 过程，那么风险溢价也会服从 $AR(1)$ 过程：$rp_t = \lambda\alpha_0 + \lambda\alpha_1 rp_{t-1}$。为了使问题易于处理，PS 在风险溢价 $\overline{rp}$ 均值附近对 RVF 进行线性化：

$$\frac{\partial[\ln P_t]}{\partial[\ln(h_t)]} = \frac{\overline{rp}}{[1 + \bar{r} + \overline{rp} - \alpha_1(1 + g)]} \tag{10}$$

因此，P_t 对波动率 h_t 变化的反应增加了持续性 α_1。PS 以一个特殊月份中 S&P 综合指数的平均每日变化为基础来计算月度股票收益的方差，即无条件波动率。因此，对 t 月份而言，

$$h_t = \sum_{i=1}^{m} s^2_{it}/m \tag{11}$$

其中，s_{it} 为 t 月份股票指数的每日变化，m 是该月的交易天数。例如，模型给出了 α_1 的取值范围大约是 0.6～0.7。他们也根据期权价格来估计隐含波动率，这就是事前波动率或前瞻性波动率，他们发现波动持续性很弱。PS 指出波动率 50% 的增加只会使股票价格下降 0.7%（他们能发现的最大效应是下降 11%）。因此，在 PS 框架下，波动持续性太低以至于无法解释所观察到的股票价格的剧烈运动。

GARCH－M 模型

Chou(1988)在 GARCH(1，1)模型的基础上使用条件方差模型来重复 PS 的研究。这就避免了 PS 两步估计方法所导致的参数不一致估计（Pagan 和 Ullah(1988)）。Chou 假设市场投资组合的预期收益由 CAPM 模型给出：

$$R_{t+1} - r_t = \lambda h_{t+1} + \varepsilon_{t+1} \tag{12}$$

其中，R_{t+1} 是市场投资组合的收益，r_t 为无风险利率，h_{t+1} 是条件方差：

$$E_t(R_{t+1}-E_tR_{t+1})^2=E_t(\varepsilon_{t+1}^2)=h_{t+1} \tag{13}$$

其中，$\varepsilon_{t+1}|\Omega_t\sim N(0,h_{t+1})$。根据 CAPM 模型，预期超额市场收益直接随条件方差而变化：大的预测误差（即更多的风险）需要用更高的预期收益来补偿。GARCH(1,1)模型为

$$h_{t+1}=\alpha_0+\alpha_1\varepsilon_t^2+\alpha_2h_t \tag{14}$$

GARCH(1,1)模型是分布二阶矩阵的一种适应性预期。$t+1$ 时刻条件方差的预测值为 $h_{t+1}=\alpha_0+(\alpha_1+\alpha_2)h_t$。进行递归替代并使用迭代期望法则就可以得到未来各期$(t+s)$的条件方差

$$h_{t+s}=\frac{\alpha_0[1-(\alpha_1+\alpha_2)^{s+1}]}{1-(\alpha_1+\alpha_2)}+(\alpha_1+\alpha_2)^sh_t \tag{15}$$

α_i 存在非负约束，所以条件方差总是非负。如果$(\alpha_1+\alpha_2)=1$，那么当前方差 h_t 的变化与未来预期之间存在一一对应关系。如果$(\alpha_1+\alpha_2)<1$，那么 h_t 对 h_{t+s}的影响按指数方式衰减。因此，$(\alpha_1+\alpha_2)$是条件方差持续性的一个度量指标。如果$(\alpha_1+\alpha_2)\geqslant 1$，那么无条件方差 $\alpha_0/[1-(\alpha_1+\alpha_2)]$没有定义（条件方差为非平稳（发散）序列）。方程（12）和（14）合在一起就构成了“GARCH－in－Mean 模型”或 GARCH－M 模型。Chou（1988）使用极大似然法同时对这两个方程进行了估计。数据是 1962—1985 期间（1225 个观察值）NYSE 价值加权股票价格指数（红利要进行再投资）的每周收益（用上周二和本周二的收盘价来进行计算）数据。

Chou 发现，风险的市场价格 λ 在各个子时期上的估计值在统计上没法很好地确定，取值边界在统计上不显著。然而，λ 的合理点估计值范围为 3～6（PS 得到的值为 3.5，Merton（1973）得到的值为 3.2）。$(\alpha_1+\alpha_2)$的值在各个子时期中非常稳定，大致为 0.98，这意味着持续性很强。根据前面的讨论可知，所观察到的股票价格急剧下降可以用 RVF 来进行解释。事实上，当$(\alpha_1+\alpha_2)=1$ 时（Chou 发现这在统计学上很容易让人接受），股票价格出现剧烈运动，弹性 $d(\ln P_t)/d(\ln h_t)$高达（负）60。

Chou 分别使用 $N=5,20,50$ 和 250 个交易日中的不同“期限”收益来对 GARCH(1,1)进行了重新估计，结果发现$(\alpha_1+\alpha_2)$的估计值非常稳定，在上述各种情形中大致为 0.95。为了进行比较，Chou 使用不同的 N 值来计算方差的 PS 指标（参见方程（11）），然后针对不同的 N 来对 $h_{N,t+1}=\alpha_0+\alpha_1h_{Nt}+v_{t+1}$进行估计，结果发现持续性 α_1 存在较大的变化，从 $N=5$ 天的 0 增加到 20 天的 0.6 和 250 天的 0.9。这意味着 PS 方法可能无法正确地捕获实际的持续程度。

作为上述情形的一个反例，考虑 Lamoureux 和 Lastrapes（1990a）对 Chou 模型进行细微修改后的模型。他们假设条件波动率要受历史预测误差（GARCH 模型）和交易量 VOL（即交易日中买/卖指令的数量）的影响：

$$h_{t+1}=\alpha_0+\alpha_1\varepsilon_t^2+\alpha_2h_{Nt}+\gamma\mathrm{VOL}_t$$

他们对每日收益进行了建模，认为预期收益为常数 μ 的假设是现实可行的：

$$R_{t+1}=\mu+\varepsilon_{t+1}\qquad\varepsilon_{t+1}\sim N(0,h_{t+1})$$

他们使用 20 家交易活跃公司一年的每日数据（1981 年或 1982 年）来对上述方程进行了估计。当将 γ 设定为零时，他们一般会得到与 Chou 相似的结果，GARCH 效应强且具有持续性（也就是说，$(\alpha_1+\alpha_2)\approx 0.8$ 到 0.95）。在将 VOL_t 添加进来后他们发现 $\alpha_1=\alpha_2=0$，但 $\gamma\neq 0$ 且残差不服从正态分布。因此，条件波动率 h_{t+1}不是由历史预测误差而是由交易量来决定的（也就是说，交易量的持续性能解释 h_{t+1}的持续性）。他们将交易量 VOL 解释为度量信息达到的一个指标，并据此猜测：其他研究中的 GARCH 效应一般都是新信息达到后实际持续性的一种度量（参见下面对市场微观结构的讨论）。因此，就这个数据集合而言，Chou 的模型对是否将 VOL_t 引入到 GARCH 过程中非常敏感。然而，Brook（1998）却认为市场交易量不会改善简单 GARCH 模型中波动率的样本外预测效果。下一节将考察 GARCH－M 模型对设定变化的敏感性和 VOL 对市场微观结构模型的重要性。

CAPM 模型与红利

Attanasio 和 Wadhwani（1990）的研究以实证观察作为开始，根据前面的工作我们知道，股票

市场综合指数（假设它可以用来代替市场投资组合）的预期超额收益取决于以前各期的红利价格比。当预期超额收益保持不变时，后者是对有效市场假说（EMH）的违背。这一领域的前期工作经常假设风险溢价为常数，但有时也将红利收益的出现解释为时变风险溢价的象征。Attanasio 和 Wadhwani 的建议是，如果我们直接对时变风险溢价进行建模，那么我们会发现红利收益（ $=Z_t$ ）不会影响预期收益。如果是这样的话，那么这是对 CAPM 模型的支持：

$$(R_{m,t+1}-r_t)=\lambda h_t+\delta Z_t+\varepsilon_{t+1}$$

我们在这里预期 $\lambda>0$ 且 $\delta=0$。假设时变条件方差可以由 GARCH（1，2）模型（增加红利收益变量）来决定：

$$h_{t+1}=\alpha_0+\alpha_1 h_t+\alpha_2\varepsilon_t^2+\alpha_2\varepsilon_{t-1}^2+\pi Z_t$$

α_1 和 π 在这里存在非负约束。因此，红利收益会影响条件方差，这并没有违背 CAPM 模型。使用 1953 年 1 月到 1988 年 11 月的美国股票价格指数的月度数据所得到的代表性结果为（Attanasio 和 Wadhwani（1990），表 2）：

$$\begin{aligned}(R_{m,t+1}-r_t)&=-0.035+0.55Z_t-4.05r_t+22.3h_{t+1}\\&\qquad(0.025)\,(0.39)\quad(1.05)\quad(11.3)\end{aligned}\tag{16}$$

$$1953(1)\sim1988(1),\ R^2=0.059,\ (\,.\,)=\text{标准差}$$

$$\begin{aligned}h_{t+1}&=\alpha_0+0.015\varepsilon_t^2+0.022\varepsilon_{t-1}^2+0.87h_t+0.053Z_t\\&\qquad(0.03)\quad(0.04)\qquad(0.06)\,(0.024)\end{aligned}\tag{17}$$

在 CAPM 方程（16）中，条件方差在统计上显著，红利收益 Z_t 不显著，结论支持 CAPM 模型。我们在方程（16）中也报告了“当 Attanasio – Wadhwani 将短期利率包含进来后的结果”，短期利率在统计上不显著，CAPM 模型遭到了拒绝。在 GARCH 方程（17）中，红利收益 Z_t 对条件方差的效应在统计上显著，这可以解释为什么那些假设风险溢价为常数的研究者会发现 Z_t 在 CAPM 收益方程中显著。注意，因为 $(\alpha_1+\alpha_2+\alpha_3)=0.91$，所以条件方差存在很强的持续性。

噪声交易者风险与序列相关

正反馈（动量）交易者在价格上升后进行购买，而在价格下跌后卖出（例如“止损”指令或“投资组合保险”策略的使用）。这就产生了收益正的序列相关性。负反馈交易者执行相反的策略：他们“买低”“卖高”。因此，如果这些交易者主导着市场，那么价格下降之后就会出现价格上升。（这对那些以相同权重来持有各种资产的投资者而言是真实的，因为资产的价格下降会导致这种资产在投资组合中的份额下降，从而产生对这种资产的额外购买，价格也会出现上升。）噪声交易者的股票需求 N_t（占股票市场总值的比例）可以表示为

$$N_t=\gamma R_{t-1}\tag{18}$$

$\gamma>0$ 表示正反馈投资者，$\gamma<0$ 表示负反馈投资者，R_{t-1} 是前期的持有期收益。我们假设精明投资者的股票需求由一个（简单的）均值－方差模型来表示：

$$S_t=(E_tR_{t+1}-\mu)/\psi_t\tag{19}$$

S_t 在这里表示精明投资者的股票持有比例，μ 是预期收益率（精明交易者所要求的预期收益率为 0），ψ 是股票感知风险的度量指标。我们假设 ψ_t 是关于股票价格条件方差 h_t 的一个为正的函数：$\psi=\psi(h_t)$。因此，精明交易者持有的股票越多，预期收益就越高，股票风险就越小。如果精明交易者持有所有的股票，那么 $S_t=1$，重新排列方程（19）就得 CAPM 模型，超额收益 $E_tR_{t-1}-\mu$ 取决于风险溢价，风险溢价在这里与股票价格的条件方差成比例：$\psi_t=\psi(h_t)$。市场均衡要求对所有的股票而言下式成立：

$$S_t+N_t=1\tag{20}$$

把方程（18）和（19）代入方程（20），进行重新整理并使用 RE 假设就可以得到：

$$R_t=\mu+\psi(h_t)-\gamma\psi(h_t)R_{t-1}+\varepsilon_t\tag{21}$$

因此，在一个存在精明交易者和噪声交易者的市场上，R_t 的序列相关性取决于噪声交易者的类型。有点自相矛盾的是，一个正反馈交易者（即 $\gamma>0$）导致了 R_t 负的序列相关性（对任意给定的

常数风险水平 h_t 而言)，反之亦然。方程(21)中 $\psi(h_t)$ 的线性表示是

$$R_t = \mu + \theta h_t + (\gamma_0 + \gamma_1 h_t) R_{t-1} + \varepsilon_t \tag{22}$$

当风险水平为常数时，反馈交易者的直接影响可以用 γ_0 的符号来表示。假设 γ_0 为正(即 R_t 存在正的序列相关性)但 γ_1 为负。随着风险 h_t 的增加，R_{t-1} 的系数(即 $\gamma_0 + \gamma_1 h_t$)会改变符号，股票收益的序列相关性由正变为负。这就说明，随着波动率的增加，市场更加受正反馈交易者所主导，他们与精明交易者之间的交互作用导致了收益的负序列相关性。

Sentana 和 Wadhwani(1992)使用 1855—1988 年期间美国每日数据和时变条件方差复杂的 GARCH 模型来对上述模型进行了估计。他们的 GARCH 模型允许无交易的天数对条件方差产生影响(French 和 Roll(1986))，然而人们在实践中却发现这在统计上并不显著。研究还发现条件方差对正负预测误差有着不同的影响(如非对称 GARCH 模型)。ε_t 一个单位的负面冲击对条件方差的影响比正的预测误差还要大。

收益正负序列相关性之间的转换点是 $h_t > (-\gamma_0/\gamma_1)$，他们发现 $\gamma_0 = 0.09$，$\gamma_1 = -0.01$，转换点为 $h_t > 5.8$。因此，当波动率很低时，极短时期中的股票收益表现出正的序列相关性，但是，当波动率很高时，股票收益却表现出负的序列相关性。因此，这个模型为“正负反馈交易者的相对影响可能会随着风险程度的变化而变化”的观点提供了统计支持(但没有解释为什么会发生这种情形)。与股票指数收益的结论相似，Sentana - Wadhwani 也发现条件方差具有较强的持续性(GARCH 模型参数的系数和接近于 1)。在实证结果中，θ 与 0 在统计上并没有显著差异，所以波动率对股票平均收益的影响只通过非线性变量 $\gamma_1 h_t R_{t-1}$ 来起作用。

Lee、Jiang 和 Indro(2002)使用 1973—1995 年期间美国的周度数据来考察“投资者情绪指标的变化”是否对(S&P500、NASDAQ 或 DJIA 指数的)预期收益方程产生直接影响，结果也发现了预期收益中的 GARCH - M 效应。他们也允许情绪指标的变化变量直接出现在 GARCH 波动率模型中，结果发现熊市情绪的增加会导致波动率的上升，对 NASDAQ 而言结论最为显著，由预期收益回归方程中的波动率变量可知，这会导致预期收益的下降。情绪指标的变化也会对预期收益产生直接影响。这些结论与噪声交易者影响预期收益和波动率的结论基本一致(与 Delong 等的模型一样)。

债券市场

在一篇开创性研究中，Lilien 和 Robins(1987)使用一个简单的 GARCH - M 模型来确定美国长期国债相对于短期国债的超额持有期收益(HPR)。人们发现，超额 HPR 正向地依赖于条件方差(系数 = 0.867，t 统计量为 5.2)，ARCH 波动过程在统计上显著(尽管 1.6 的系数之和比 1 要大一些)。然后，Engle 等人将收益利差($R_t - r_t$)纳入到超额收益方程之中并发现这个变量在统计上显著，因此 CAPM 模型遭到了违背。

尽管 Engle 等人的结果似乎说明条件方差对预期均衡收益有着很强的影响，Tzavalis 和 Wickens(1995)却认为结论对数据时期的选择非常敏感，当我们将 1979—1982 年的利率极度波动时期(美国的货币政策目标开始生效)考虑进来后，结果就变得尤其敏感。从宽泛意义上讲，Tzavalis 和 Wickens(使用月度数据)使用 GARCH(1,1)模型对 Engle 等人的结论进行了重新考察，但新增了一个虚拟变量 DV_t：在 1979(10)到 1982(9)的各个月份中虚拟变量取值为 1，否则取值为 0。他们发现，在增加虚拟变量 DV_t 后，波动持续性有所减小，也就是说，GARCH 过程中的($\alpha_1 + \alpha_2$)变为 0.3 而不是 0.9(虚拟变量高度显著且为正)。此外，预期 HPR 不再受条件方差影响，所以预期收益不存在 GARCH - M 效应。当然，虚拟变量只是增加了 1979—1982 年期间波动率的平均水平，为什么会出现这种外生的位移，我们仍然无法进行解释。因此，根据直观的经济结论，人们也会支持无虚拟变量的模型并持有如下观点：在波动率的水平值很高时，波动持续性也会很强。阈值效应也可能存在，因此，波动持续性在高波动时期要强一些，也会对均衡预期收益产生影响。反之，波动持续性在低波动率时期要弱一些，相对较小的条件方差值不会对均衡收益产生很大的影响。从直觉来看，上述结论似乎很合理，但是，一个显然的事实是，我们需要对这

种非对称的阈值效应进行更为深入的研究。

GARCH 模型：一些变体

简单模型的变体比早餐种类还要多，许多变体将不对称引入到波动率过程之中，背后的经济直觉是：股票收益出现 $x\%$ 的突然下跌所导致的波动程度比股票出现等量突然上升所导致的波动程度要高。通过在 GARCH 方程引入“新的”经济变量（如杠杆率，也就是说，负债权益比）或对 GARCH 过程本身进行修正就可以将这种不对称效应引入到模型之中。例如，Glosten、Jagannathan 和 Runkle（GJR）（1993）模型增加了一个指标变量：

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 h_{t-1} + \gamma(\varepsilon_{t-1}^2 D_{t-1}) \tag{23}$$

其中，

$$D_{t-1} = \begin{cases} 1 & \text{如果 } \varepsilon_{t-1} < 0 \\ 0 & \text{其他情形} \end{cases}$$

一个更为复杂的不对称模型是 Nelson（1991）的 EGARCH 模型，

$$\ln h_t = \alpha_0 + \alpha_1 \ln h_{t-1} + \gamma \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \left[\frac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}}} - \sqrt{\frac{2}{\pi}}\right] \tag{24}$$

h_t 在这里永远非负（即使参数可以为负），模型适用于 ε_t 很一般的分布，尽管人们在实践中经常使用条件正态分布。如果 $\gamma < 0$，那么收益负的意外冲击所导致的条件波动率变化比正的意外冲击所导致的变化大。这些模型用于每日、每周、每月股票收益的研究一般都会发现不对称效应的实证证据，但这些效应不是特别的强。

预测效果

各种 GARCH 模型与其他波动率模型相比的效果如何呢？Day 和 Lewis（1992）使用 S&P100 指数在 1983.3 ~ 1989.12 期间的周度数据来对 GARCH - M 模型进行了估计：

$$(R_{t+1} - r_t) = \lambda_0 + \lambda \sqrt{h_t} + \varepsilon_{t+1} \tag{25}$$

他们考虑了 GARCH(1,1) 和 EGARCH 过程，两者的样本内结果之间几乎没有差异，两者的“均值效应”在统计上不显著（即 $\lambda = 0$）。然而，隐含波动率序列（来自于期权价格）确实增加了 GARCH 和 EGARCH 模型的解释力，尽管 GARCH 和 EGARCH 模型本身在样本中也表现得很好。

更为严格的模型检验是考察样本外预测效果。使用滚动窗口就可以每周对模型进行更新，然后用这些模型来预测一周后的波动率。他们用两个事后指标来刻画波动率：每周收益的平方（$V_t^w = R_{w,t}^2$）和周内每日收益的平方（$V_t^d = \sum_{i=1}^{5}(R_i - \mu)^2$，$R^2$ = 每日收益，μ = 每日收益的均值）。这些每周波动率的事后指标在提前一周的预测回归中可以用作因变量。其他预测结果可以由两个 *GARCH* 模型或隐含波动率（来自于期权）模型或波动率的简单历史移动平均（即 $V_t^{ma} = \sum_i^m = 1\ R_{i-1}^2/m$）来给出。结果很令人失望，因为隐含波动率序列和两个 GARCH 模型对（两个）事后波动序列 V_t^w 和 V_t^d 在统计上没有解释力（$R^2 \approx 0.025$）。事实上，历史波动率 V_t^{ma} 是有助于预测两个事后波动率序列的唯一变量。R^2 在这里仍然较低，约为 0.038，因此我们很难用其他方法来对波动率进行预测。在波动率高度变化的情形中这并不令人感到很奇怪。然而，这可以解释为什么在用 VCV 方法对每日 Var 进行预测时金融机构倾向于使用像 EWMA 那样的简单预测方程。

也许值得提及的是随机波动率模型，这种模型在条件方差方程中增加了一个随机项：

$$y_{t+1} = \mu + \sqrt{h_t}\varepsilon_{t+1} \quad \varepsilon_{t+1} \sim N(0,1) \tag{26}$$

$$\ln h_{t+1} = \alpha_0 + \alpha_1 \ln h_t + \sigma_u u_{t+1}^2 \tag{27}$$

其中，$u_t \sim N(0,1)$。随机波动率模型和 MCS 结合在一起就可以用来对期权进行定价，这里假设波动率（和标的资产的价格）是随机变量。随机波动率模型可以用 Kalman 滤波方法来进行估计，但我们不对此进行更为深入的讨论（参见 Cuthbertson 和 Nitzsche（2001b），Ruiz 和 Shephard（1994））。

29.3 多变量 GARCH 模型

我们在前面的章节中提到过使用多变量 GARCH 模型来对 SDF 模型进行检验。最简单的多变量 GARCH 模型由两个资产收益方程组成，当然也涉及两个条件方差和一个条件协方差。两种资产平均收益最为简单的方程为

$$R_{t+1} - r_t = \mu + \varepsilon_{t+1}$$

$R_{t+1} - r_t$、μ 和 ε_{t+1} 在这里都是(2 × 1)向量。现在转到 *GARCH* 过程(2 × 1)误差项 $\varepsilon_{t+1} = \{\varepsilon_{1,t+1}, \varepsilon_{2,t+1}\}'$的条件协方差上去，即使是无约束模型也存在很多待估参数。因此，就实践而言，人们施加一些特殊条件的目的是为了减少待估参数的数量。符号“VECH”在多变量 GARCH 模型中经常使用(Bollerslev、Engle 和 Wooldridge(1988))，对两种资产的情形而言我们有(在 GARCH 过程中只有一阶滞后值)：

$$\begin{aligned} &\mathrm{VECH}(H_{t+1}) = C + A[\mathit{VECH}(\textstyle\sum_t \sum_t')] + B[\mathrm{VECH}(H_t)] \\ &\textstyle\sum_t \mid \Omega_t \sim N(0, H_t) \end{aligned} \tag{28}$$

其中，

$$H_t = \begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{21,t} & h_{22,t} \end{bmatrix}$$

$$\sum{}_t = \begin{bmatrix} \varepsilon_{11,t} & \varepsilon_{12,t} \\ \varepsilon_{21,t} & \varepsilon_{22,t} \end{bmatrix}$$

VECH 算子将矩阵的上(或下)三角部分变换成一个简单的列向量，例如，

$$\mathrm{VECH}(H_t) = [h_{11}, h_{22}, h_{12}]_t'$$

$$\begin{aligned} \mathrm{VECH}(\textstyle\sum_t \sum_t') &= \mathit{VECH}\begin{bmatrix} \varepsilon_{11}^2 + \varepsilon_{12}^2 & \varepsilon_{11}\varepsilon_{21} + \varepsilon_{12}\varepsilon_{22} \\ \varepsilon_{21}\varepsilon_{11} + \varepsilon_{22}\varepsilon_{12} & \varepsilon_{21}^2 + \varepsilon_{22}^2 \end{bmatrix}_t \\ &= [\varepsilon_{11}^2 + \varepsilon_{12}^2, \varepsilon_{21}^2 + \varepsilon_{22}^2, \varepsilon_{11}\varepsilon_{21} + \varepsilon_{12}\varepsilon_{22}]_t' \end{aligned}$$

VECH(H_t)和 VECH($\sum_t \sum_t'$)都是(3 × 1)向量，A 和 B 都是(3 ×3)系数矩阵，C 是一个(3 × 1)列向量。如果没有限制条件，那么 VECH 模型有 h_{11} 这一项(举个例子)，它取决于(h_{11}, h_{22}, h_{12})中每个分量的滞后值和 VECH($\sum \sum'$)中三个不同的误差项，方程共有七个参数(包括常数 c_{11})。例如，$h_{11,t+1}$所对应的第一个方程为

$$\begin{aligned} h_{11,t+1} = {} & c_{11} + a_{11}(\varepsilon_{11,t}^2 + \varepsilon_{12,t}^2) + a_{12}(\varepsilon_{21,t}^2 + \varepsilon_{22,t}^2) + a_{13}(\varepsilon_{11,t}\varepsilon_{21,t} + \varepsilon_{12,t}\varepsilon_{22,t}) \\ & + b_{11}h_{11,t} + b_{12}h_{22,t} + b_{13}h_{12,t} \end{aligned}$$

因此，由两种资产组成的体系在(h_{11}, h_{22}, h_{12})所对应的三个 GARCH 型方程中有 3 ×7 = 21 个参数，许多参数都需要进行估计(即使是针对这个两资产情形)。为了简化模型，我们可以将 A 和 B 限制为对角矩阵，从而每个条件方差(协方差)变为

$$h_{ij,t} = \alpha_{ij} + \beta_{ij}(\varepsilon_i \varepsilon_j)_{t-1} + \gamma_{ij} h_{ij,t-1} \tag{29}$$

这就将(两资产模型的)待估参数的数量减少到 9 个。我们必须确信所估计的参数能给出一个半正定矩阵，但这没法进行保证。“正定”意味着对角线上的元素为正且非对角部分对称——这是方差协方差的一个合意性质。半正定协方差矩阵也意味着，对任何投资组合权重而言，投资组合的方差大于或等于 0，这是另外一个合意性质。

一个稍微不同的模型保证了 H 矩阵始终是正定矩阵，这就得到了“BEKK 模型”(参见 Engle 和 Kroner(1995))：

$$H_{t+1} = C'C + A'H_tA + B'(\textstyle\sum_t \sum_t')B \tag{30}$$

对 $n=3$ 种资产（比方说）的情形而言，C、A 和 B 都是矩阵，二次项保证了正定性。当 $n=3$ 时，常用的模型是 SDF 模型或 GARCH 过程，在 GARCH 过程中，h_t 和 ε_t^2 的一阶滞后值会出现（即文献中常说的 GARCH($p=1, q=1$)模型或 GARCH(1,1)模型），于是 BEKK 模型需要估计 $n(n+1)/2+(p+q)n^2(n+1)^2/4=78$ 个参数。这个模型的向量误差修正模型 VECM(1,1) – BEKK 为

$$H_{t+1}=V'V+A'(H_t-V'V)A+B'(\varepsilon_t\varepsilon_t'-V'V)B$$

这里的 V 是"长期"或无条件协方差矩阵（可以以样本均值作为开始来进行计算）。其他项表示与长期均衡的偏差。（这个方程就是我们在第 2 章所见过的误差修正模型 ECM，但这里使用了二阶矩）。当 $n=3$ 时，VECM – BEKK 模型中的待估参数有 $3n^2=27$ 个。我们可以对这个模型进行进一步限制，使得 V、A 和 B 都是下三角矩阵，进行限制后的 VECM – BEKK 模型只需对 $3n(n+1)/2=18$ 个参数进行估计。最后，如果将 $\varepsilon_{i,t+1}$ 和 $\varepsilon_{j,t+1}$ 之间的相关系数设定为常数，那么我们可以对估计过程进行简化

$$h_{ij,t+1}=\rho_{j,t+1}[h_{ii,t+1}h_{jj,t+1}]^{1/2}$$

$$h_{ii,t+1}=c_i+\alpha_i h_{ii,t}+\beta_i\varepsilon_{it}^2$$

这里有 $3n+n(n-1)/2=12$ 个待估参数。但是，正如我们试图对时变方差和协方差进行建模那样，这是多变量 GARCH 过程一个带限制性的版本（如果方差具有时变性，那么这个模型只允许协方差随着时间进行变化（$\sigma_{ij,t}=\rho_{ij}\sigma_{i,t}\sigma_{j,t}$））。显然，人们应在多变量 *GARCH* 模型中的协方差数目与 *GARCH* 模型的使用弹性之间进行权衡，前面各章对股票、债券和外汇收益进行建模时已经注意到了这一点。

期货套期保值

我们以股指期货的动态套期保值说明多变量 GARCH 模型的用途。很容易证明（Cuthbertson 和 Nitzche(2001b)），使分散良好的投资组合的套期保值误差达到最小的最优股指期货合约数目为

$$N_{f,t}=-\left(\frac{\mathrm{TVS}_0}{zF_0}\right)\beta_t \quad \text{（即买空 } N_{f,t} \text{份期货合约）} \tag{31}$$

其中，TVS_0 = 股票组合的现金价值，F_0 = 当前的期货价格，z = 一个指数点所对应的价值（例如，对 S&P500 股指期货合约而言，一个点的价值为 250 美元），且

$$\beta_t=h_{12,t+1}/h_{22,t+1} \tag{32}$$

h_{22} 是期货价格变化的条件方差，h_{12} 是投资组合中股票价格变化与期货价格变化之间的条件协方差。如果套期保值是"完全"套期保值，那么股票期货组合的价值变化量为 0。最简单的套期保值比率是静态套期保值比率，β 在这里是固定的且由 $\Delta S_t/S_{t-1}$ 对 $\Delta F_t/F_{t-1}$ 的 *OLS* 回归系数来得到。此外，我们可以使用滚动窗口来对 $(h_{11},h_{22},h_{12})_t$ 的多变量 GARCH 模型进行估计并进行单步预测，然后在每一期中使用最小方差套期保值比率 β_t。Brooks、Henry 和 Persand(2002)进行每日套期保值的结果是，静态套期保值使得样本外套期保值误差减小了约 90%。使用模型还会使套期保值误差再减少 2%，但增加 GJR 不对称效应无法对结果进行改进（Brooks、Henry 和 Persand(2002)，也可以参见 Cumby、Huizinga 和 Obstfeld(1983)，Franses 和 vanDijk(1996)）。这些研究说明 GARCH 类模型使用动态套期保值后的效果是对传统静态（OLS）套期保值比率的一种改进。

债券收益

多变量 GARCH 模型的使用经常涉及大量的参数估计，且估计是很强的非线性估计，这会产生一些（收敛）问题。由于这些技术困难，特别是在一个有限数据集合给定的条件下，研究人员通常会对资产的种类数进行限制，也会对 GARCH 过程的参数施加某种限制。

我们接下来要讨论的债券市场理论模型将考虑各种期限的债券，也会涉及时变风险溢价。基本模型是 CAPM 模型，因此，时变溢价可以通过时变协方差和多变量 GARCH 模型来进行建模。我们以收益方程作为讨论的开始。对市场投资组合而言，超额 HPR（$=y_{m,t+1}$）与风险的市场

价格 λ 和市场收益的条件方差成比例：

$$(y_{m,t+1}) = \lambda \operatorname{Var}_t(y_{m,t+1}) + \varepsilon_{m,t+1} \tag{33}$$

我们在这里假设 $\varepsilon_{m,t+1} \sim N(0, h_{m,t+1})$。对期限为 n 的债券而言，超额 HPR（$= y_{n,t+1}$）取决于条件协方差：

$$y_{n,t+1} = \lambda \operatorname{Cov}_t(y_{n,t+1}, y_{m,t+1}) + \varepsilon_{n,t+1} \tag{34}$$

其中，

$$\operatorname{Cov}_t(y_{n,t+1}, y_{m,t+1}) = E(\varepsilon_{n,t+1}, \varepsilon_{m,t+1}) = h_{mn,t+1} \tag{35}$$

因此，我们可以将（34）写成

$$y_{n,t+1} = \lambda h_{mn,t+1} + \varepsilon_{n,t+1} \tag{36}$$

对 n 期债券而言，条件方差和（与市场收益之间的）条件协方差可以用 GARCH（1，1）方程来进行建模：

$$h_{n,t+1} = \alpha_0 + \alpha_1 h_{n,t} + \alpha_2 \varepsilon_{n,t}^2 \tag{37a}$$

$$h_{mn,t+1} = \delta_0 + \delta_1 h_{nm,t} + \delta_2 \varepsilon_{n,t} \varepsilon_{m,t} \tag{37b}$$

市场的 GARCH 条件方差为

$$h_{m,t+1} = \gamma_0 + \gamma_1 h_{m,t} + \gamma_2 \varepsilon_{m,t}^2 \tag{38}$$

如果我们只有一种 n 期债券和一个市场投资组合，那么在预期收益方程中我们就有 9 个 GARCH 参数需要估计，还要外加 λ。当我们将不同期限的各种债券添加进来后，GARCH 模型的参数数量急剧上升。因此，在对各种期限债券的上述模型进行估计后，Hall、Miles 和 Taylor（1989）假设每个期限的债券对应的 GARCH 参数数目相同。此外，尽管使用的是好几个国家的债券数据，他们还是假设市场投资组合只包含一种国内债券（即不包含外国债券）——这也可以减少 GARCH 协方差项的数量。他们使用的数据是月度数据。对债券市场投资组合而言，有证据表明方差中存在 GARCH 效应，因为 γ_1 和 γ_2 不等于零。然而，条件方差似乎无法（至少在统计上）解释超额市场 HPR$y_{m,t+1}$的大部分变化，因为在方程（33）中，只是对日本和法国而言 λ 才与 0 在统计上显著不同（对英国、加拿大、美国和德国而言我们可以接受）。*Hall* 等人也发现，就多个国家（如英国、日本、美国和德国）而言，条件协方差项对 n（$n = 1 \sim 3, 3 \sim 5, 5 \sim 7, 7 \sim 9, 10 \sim 15$ 及大于 15 年）期债券的 HPR 没有影响。

Hall 等人也考察了 t 时刻的信息（超额 HPR 的滞后值和收益利差）是否会影响超额 HPR。在绝大多数情形中，他们都拒绝了“这两个变量在统计上都显著”的零假设。然而，在尤其短的市场上（如 1 ~ 3 年期债券），他们发现时变方差 h_t 无论在数量上还是在统计上都比时变协方差 $h_{nm,t}$重要，因此 CAPM 模型遭到了拒绝。

对 Hall 等针对不同期限和不同国家所发现的复杂结论而言存在多种可能的解释。第一，让市场投资组合包含所有国内债券（不包括国外债券，也不包括国内和国外权益资产）。第二，由于模型存在大量的待估参数，所以 Hall 等人假设方差和协方差不同 GARCH 方程中的参数数目相等（即 $\alpha_i = \delta_i = \gamma_i (i = 1, 2)$）。这就保留了更大的自由度，对高度非线性极大似然函数进行最大化就可以减轻困难的难度，但也会产生一些无效约束。

尽管 Hall 等人在研究中进行了一些限制性假设，结果还是发现债券市场确实存在时变风险溢价现象（呈现出持续性），但这很难获得实证研究的支持，对不同期限和不同国家的债券而言这种效应也不一致。他们的研究确实强调了使用多变量 GARCH - M 方法来对具有时变方差和协方差的 CAPM 进行建模时的困难。待估参数的数目和受限（也可能质量有点糟糕）数据集合上非线性最大化程序意味着任何结果都可能存在很大的误差，对较小的设定变化而言这些结果可能很不稳健。

交互作用：股票市场与债券市场

在 CAP 模型中，所有投资者都以市场价值为权重来持有市场投资组合。然而，我们知道持有大约 25 种资产就可以实现分散化的绝大部分收益。在交易成本、信息收集成本、监督成本及

对即将到期的资产进行套期保值的资金需求给定的条件下，对一些个体和机构而言持有市场投资组合的一个子集是有意义的。因此，投资者关心的可能是如何在一个资产群体之间进行选择，相对而言他们不是很关心如何在一个特殊的“群体”内进行选择。因此，投资者可能关心如何在期限为六个月的债券、一组长期债券和一组股票之间进行决策。Bollerslex、Engle 和 Wooldridge（1988）在使用 CAPM 模型对这类资产的三月期 *HPR* 进行建模就使用了这个简化假设。很容易证明，如果我们用这三类资产来构造市场投资组合，那么

$$\mathrm{Cov}(y_1, y_n) = \mathrm{Cov}\left(y_1, \sum_{1}^{3} w_j y_j\right) = w_1\sigma_{11} + w_2\sigma_{12t} + w_3\sigma_{13t} \tag{39}$$

w 在这里表示市场投资组合中各种资产的价值权重。因此，我们可以将“单个”CAPM 协方差项从上述三个部分中分离出来。三种资产的超额收益方程为

$$y_{i,t+1} = \alpha_i + \lambda\left[\sum_{j=1}^{3} w_j h_{ij,t}\right] \tag{40}$$

其中，w 是已知的，λ 是风险的市场价格，根据 CAPM 模型，所有资产的值都相等。Bollerslev 等对形如（40）的三个超额 HPR 方程进行了估计：六月期债券、20 年期债券和股票市场指数的相应方程。他们使用多变量对角模型来对三个时变方差和协方差项进行建模，$(h_{11}, h_{12}, h_{22}, h_{13}, h_{23}, h_{33})$ 相应的形式为：

$$h_{ij,t+1} = \alpha_{ij} + \beta_{ij} h_{ij,t} + \gamma_{ij}\varepsilon_{it}\varepsilon_{jt} \tag{41}$$

由美国 1959（1）～1984（2）期间的数据所得出的结论有：

（ⅰ）因为 $\lambda = 0.499$（$s.e. = 0.16$），所以三种资产的超额持有期收益取决于时变条件协方差。

（ⅱ）条件方差和协方差都具有时变性，使用对角形式的模型可以很好地对其进行建模。

（ⅲ）短期债券的条件方差持续性最强，其次是长期债券，最后是股票（$\beta_{ii} + \gamma_{ii}$ 分别等于 0.91、0.62 和 0.55）。

我们将由 GARCH 方程所刻画的方差加入到（40）后，却发现在统计上不显著（正如我们在 CAPM 模型下所预测的那样）。然而，CAPM 模型确实无法对超额 HPR 进行“完全”解释，因为在将资产 i 的滞后 HPR 引入到方程（40）后却发现在统计上非常显著，这就拒绝了标准 CAPM 模型。

即期外汇市场上的波动率

大量研究使用 ARCH 和 GARCH 模型来对即期外汇市场的时变波动率进行建模。早期研究使用单变量模型（Engle 和 Bollerslev（1986），Diebold（1988），Baillie 和 Bollerslev（1989），Bollerslev、Chou 和 Kroner（1992））发现交易日内、每日和每周数据中存在很强的 ARCH 效应，就月度数据而言，这种效应要弱一些，对超过一个月的期限而言，ARCH 效应几乎消失。我们在 23 章中对 GARCH 模型的作用进行了讨论：条件协方差会影响外汇投机活动的预期收益。这些 GARCH 效应代表了时变风险溢价。尽管这些模型也考虑了波动率（和协方差）的持续性，但 GARCH－M 效应的统计显著性对不同货币和不同时期而言非常不稳健。

波动传染

考察波动持续性的主流文献认为市场上的消息或新信息会对另外一个市场上的波动率产生影响。将 $t-1$ 时刻的误差项 $\varepsilon^2_{j,t-1}$（对市场 $j = 1, 2, \ldots, m$ 而言）纳入到市场 $i(i \neq j)$ 波动率 h_{it} 的方程中就可以对此进行检验。Engle、Ito 和 Lin（1990）及 Ito、Engle 和 Lin（1992）使用的是“流星雨与热浪”的比喻。纽约的热天后面往往还是热天，但一般不可能出现的情形是东京会出现热天——这就是热浪型“消息”的思想。另外，纽约的一场流星雨之后几乎可以肯定的是东京会也会有一场流星雨，也就是说，“消息”会从一个市场“溢出”到另外一个市场（存在一定的时滞）。使用美元－日元汇率的日间数据，Engle、Ito 和 Lin 发现消息就像流星雨在全世界不同的市

www.55188.com最好的股票论坛理想在线证券网欢迎您!

场之间流动。

29.4 市场微观结构——外汇交易

有关市场微观结构的大量实证研究集中关注外汇交易市场，部分原因是，我们可以从外汇市场上获得高质量的高频数据。直到最近，外汇市场还是高度“分散”的。与集中交易的市场（如伦敦股票交易所、纽约证券交易所）——交易在几分钟内就公之于众——不同，外汇市场上的许多交易是交易双方（或多方）通过电话、电报或计算机网络进行的，交易也不是公开进行的。市场参与者有做市商、经纪商和顾客。全球大约有100个做市商（典型的是大银行）时刻准备按照屏幕指导价格和限定价格（限价指令）进行货币买卖（“双向”价格）。做市商在接到电话后报出买卖价格，买卖差价可以使做市商在一个交易回合之后获得一定的利润。一次交易的平均时间约为10～15分钟，这段时间中的风险是外汇报价差价的主要决定因素（Lyons（1995，1996））。Lyons（1998*b*）发现，每日交易量为10亿美元的外汇经纪商可以获得大约10万美元的利润。“顾客”不是经纪商，他们可能是其他银行、中央银行、养老基金和共同基金（包括对冲基金）中的客户。做市商可以用自己的账户进行交易，而经纪商则不行，他们通过买卖限价指令将钱存到做市商的“账户”上，通过内部报价差价（也就是说，对特殊限价指令而言的最优买卖价格）来获取利润。经纪商将市场参与者连接在一起并进行匿名交易从而收取佣金。

顾客在即期市场上的交易量相对较小（美国大约为5%），经纪商之间的直接交易约占即期交易40%，以经纪商为中介进行的交易约占55%。信用风险是这些交易存在的一个问题，经纪商会很谨慎地考虑暴露于交易对手的信用风险高低，这是为什么绝大多数外汇交易都是在经纪商之间进行的。两个主要的经纪商间外汇经纪交易系统是Reuters Dealing 2002－2和Electronic Broking Services（EBS）－Spot Dealing System。这些系统对相应银行的净信用限额进行跟踪，所以交易者知道交易是牌价交易（尽管交易对手是匿名的）。

近年来，人们对用微观结构方法来分析外汇交易市场的兴趣与日俱增。这种方法的新颖之处在于分析实际交易过程，关注指令流（order flow）和买卖价差。这种方法对外汇市场的引人之处体现在三个方面。第一，银行间市场电子交易量的增长使我们可以获得高质量的微结构数据，而基于电话的市场交易使得人们很难对交易过程进行跟踪。第二，传统模型无法预测或解释外汇变动使得新方法的更加引人注意——传统模型对外汇汇率建模的实际效果为4%的R^2，而微观市场结构模型使得R^2高达到60%（尽管用R^2值来反映对外汇运动的实际解释力不是很有意义）。第三，外汇市场的一个困惑是它所产生的巨额指令流，人们期望微结构方法能解释这一困惑以及外汇市场上的其他困惑，如远期贴水之谜。

外汇市场电子交易量的增长意味着（至少对银行间（或经纪商间）市场而言是这样），与Richard Lyons（1999）中所描述的情形相比较，市场几乎没有被承认过。因为电话交易的分散性和不透明性，外汇交易市场对交易双方而言都具有私密性，交易之前无法获得确切的买卖价格信息，主要货币几乎所有的交易都是通过两个电子交易系统（EBS和Reuters－2002）来进行的。电子交易系统能有效地对订单进行登记，市场参与者可以公布他们即将进行交易的价格和数量（限价指令），也可以接受已有的限价指令（市场指令）。因此，交易者在交易屏幕上能看到确切的市场买卖价格（交易前的透明性）、最近发生的交易数量和交易价格（交易后的透明性）。不幸的是，对最终的顾客而言这些信息仍然不可获得，因为只有经纪商才可以看见这些屏幕交易信息。因此，对公司客户、机构投资者甚至对冲基金而言，市场仍是不透明的，因为他们只能间接地获得买卖价格信息（比如通过Reuters’ FXFX系统）并通过电话进行交易。

2001年（BIS（2001）），外汇每日全球交易额达到1.2万亿，远期交易（直接外汇远期交易和外汇互换）约占交易量的65%。美元交易占全球交易的90%，因为美元充当了其他货币交易的媒介。德国马克（现在是欧元）是交易量居于第二位的货币，接下来是日元和英镑。英国市场占

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

总交易的31%，接下来是美国(16%)、日本(9%)和新加坡(6%)。做市商提供"双向拍卖"价格(即他们的买价和卖价报价)，但经纪商市场却是"单向拍卖价格"市场(只提供买价或卖价，但不会同时提供买卖价格)。

尽管交易总量从1998年开始有所下降(主要是欧元开始流通)，但外汇市场上的交易活动仍然使其他主要市场(比如说债券市场和股票市场)相形见绌。交易量是如此巨大，而许多传统外汇交易模型却无法解释汇率变动和高汇率波动率，这使很多人都想问如下问题：外汇交易本身是否能解释汇率的绝大部分变动(而不是传统视角：交易只是对新信息的一种反应)。这方面的研究是外汇市场微观结构文献中最吸引眼球的一个领域。人们从上述外汇市场微观结构视角来对买卖差价的大小、市场交易量和波动率的决定因素进行了考察。

买卖价差

经纪商间交易的低透明性可能有助于减小买卖差价，这是因为，大的(买或卖)客户指令在经纪商与客户初始交易大的不平衡信息被揭示之前就已经在众多经纪商之间流传。因此，低透明性导致了经纪商间更高水平的风险共担(在任何价格效应被感知之前)，从而买卖价差更小。Hau、Killeen和Moore(2002)认为，这一机制在1999年1月1日引入欧元之后的第一年中非常适用。"货币数量的减少"增加了欧元经纪商间货币失衡方面的信息，风险增加(也就是说，当你持有货币时价格却向相反的方向运动)使得价差变大。例如，德国马克向欧元的转换导致了平均价差的升高，与美元之间的汇率价差从3.76个基点增加到5.26个基点，与日元之间的汇率价差从5.1个基点增加到8.3个基点，与英镑之间的汇率价差从3.1个基点增加到9.2个基点(与欧元之前的时期进行比较)。根据其他微结构研究的结果，我们知道价差和交易量具有反向关系(Hau、Killeen和Moore(2002)针对欧元运行的第一年证实了这一结果)，而价差与汇率波动率成比例(Glosten和Milgrom(1985)，Admati和Pfleiderer(1998))，也与价格风险的预测值成比例(Bessembinder(1994))。

指令流

指令流(order flow)作为短期汇率变动近似决定因素的实证重要性最近被学者们所考察(Lyons(1999)，Evans和Lyons(1999))。"指令流"是买方初始指令减去卖方初始指令后的净值，其变化代表的是货币净需求的变化。一种解释是"指令流"捕获了市场参与者所能获得的私人信息或公开信息，从而对价格变化进行了更好的描述。从本质来看，这种方法假设一些交易者比其他交易者拥有更多的信息，市场通过对指令流进行监视和反应来推断这些交易者所能获得的信息。Evans和Lyons(1999)认为指令流是许多双边每日汇率变化(在4个月或更短的时期中)的重要决定因素，短期样本外预测值比随机游走模型的预测结果好。Lyons(1995)和Bessembinder(1994)建立了存货持有成本(inventory carrying co)的增加与利差扩大之间的联系。

29.5　调查数据与预期

我们使用调查数据来考察预期变化是否导致汇率变动幅度的减小。如果大家阅读过大众读物，那么你会认为外汇经纪商都是最优秀的投机者。穿着条文衬衣和时尚套装的年轻人在电视上频繁亮相，为了快速执行外汇买卖指令对着两步电话同时喊叫。显而易见的问题是，个体买卖外汇是否是基于基本信息？他们是否是"趋势追逐者"？如果是第一种情形，那么接下来的问题是，他们对外汇价格是否具有持续的影响？正如我们已经见到的那样，使用FRU的多种方法可以用来对市场效率进行检验。然而，对外汇经纪商的使用技巧及其是否导致与基础信息不相关的汇率运动进行考察的却很少。一个很有用的代表性研究是Allen和Taylor(1989)，他们对外汇市场一个非常小的部分进行了观察，对图表主义行为进行了调查。图表专家只研究市场上的

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

价格运动，将对未来的看法建立在历史价格变化基础之上，相信自己能发现历史价格运动中的一些模式，并以此来预测未来的价格运动，制定获利交易策略。当然，在有效市场中这种情形不大可能出现。

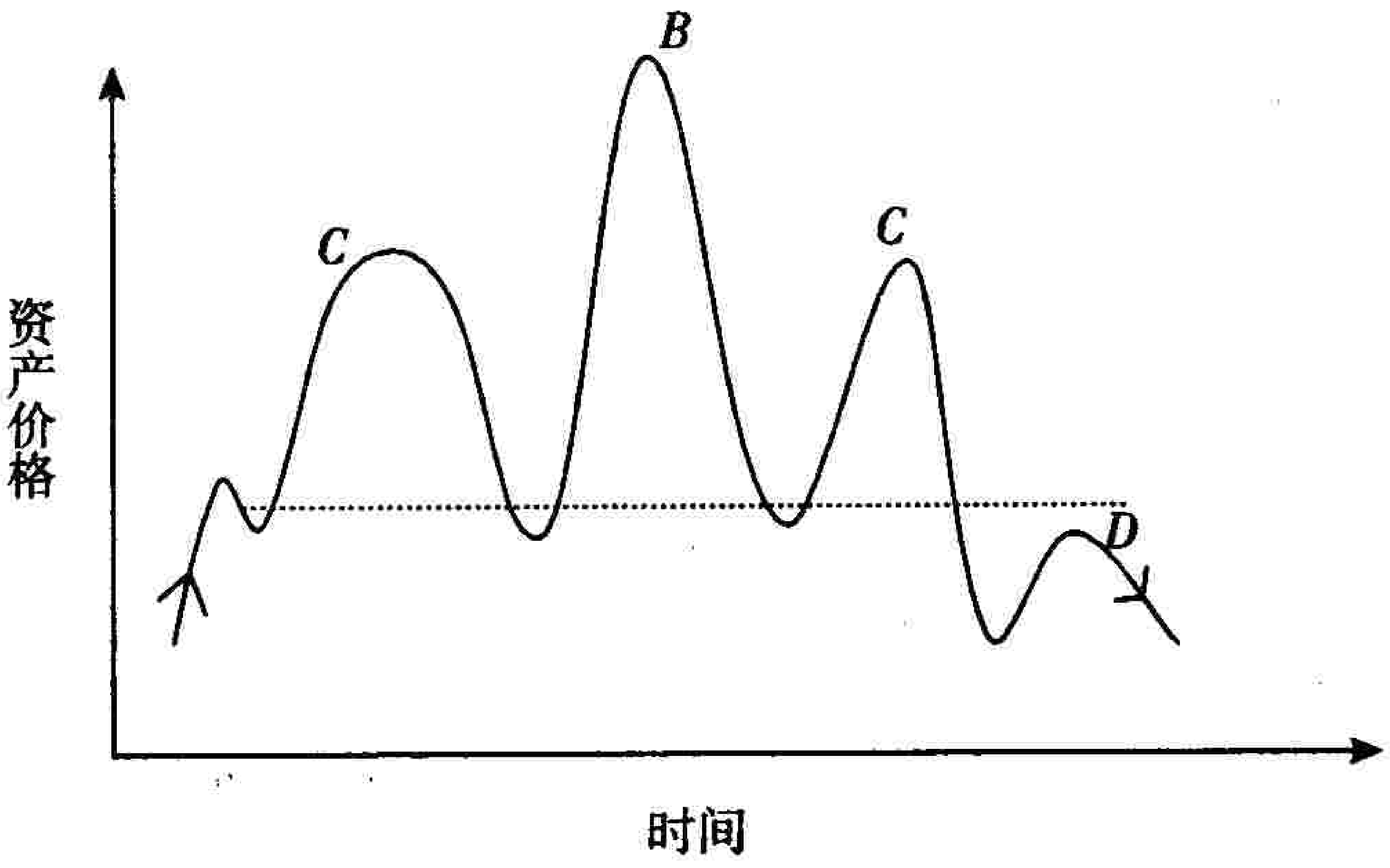


图 1　头肩图形

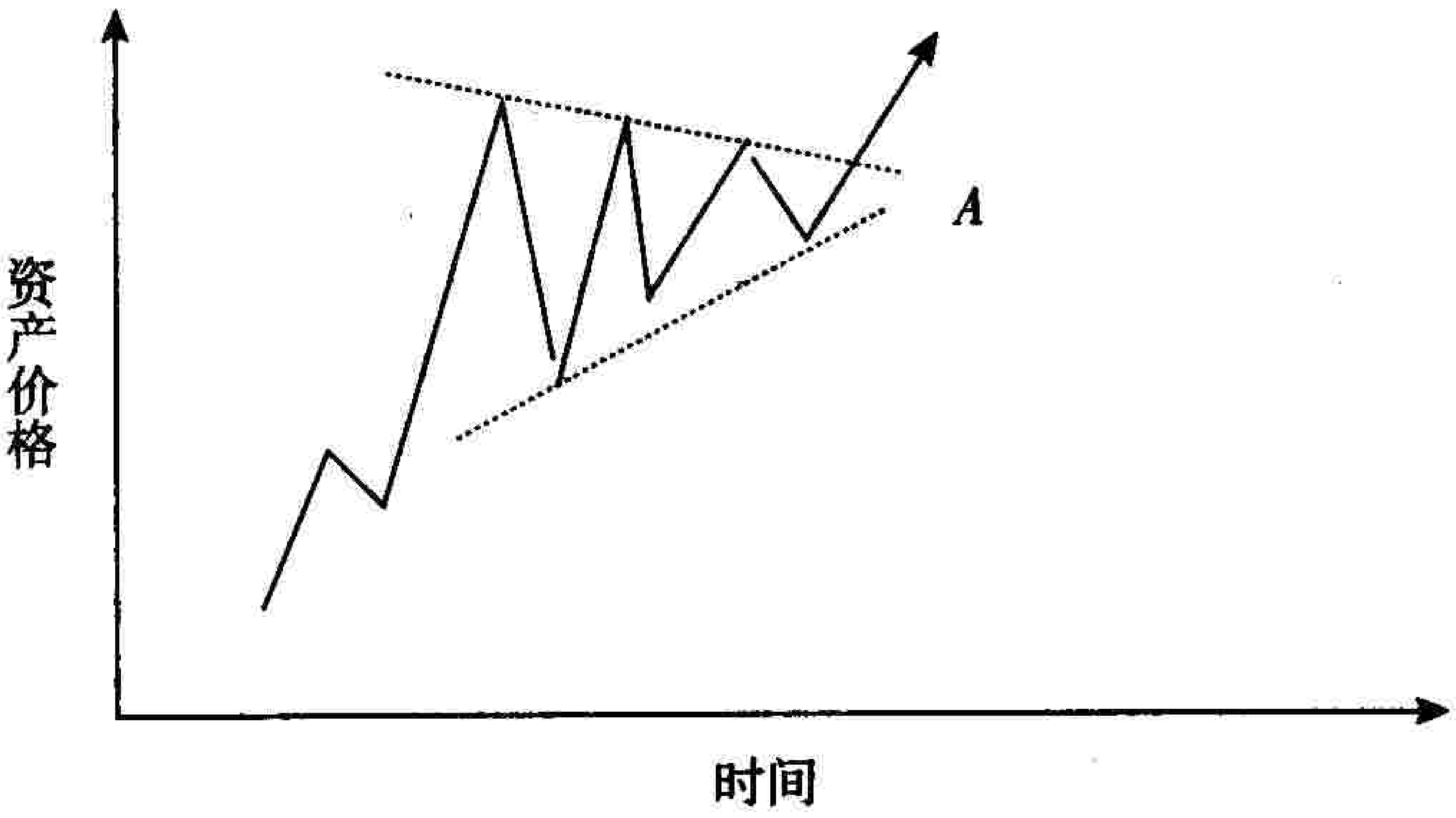


图 2　对称三角形

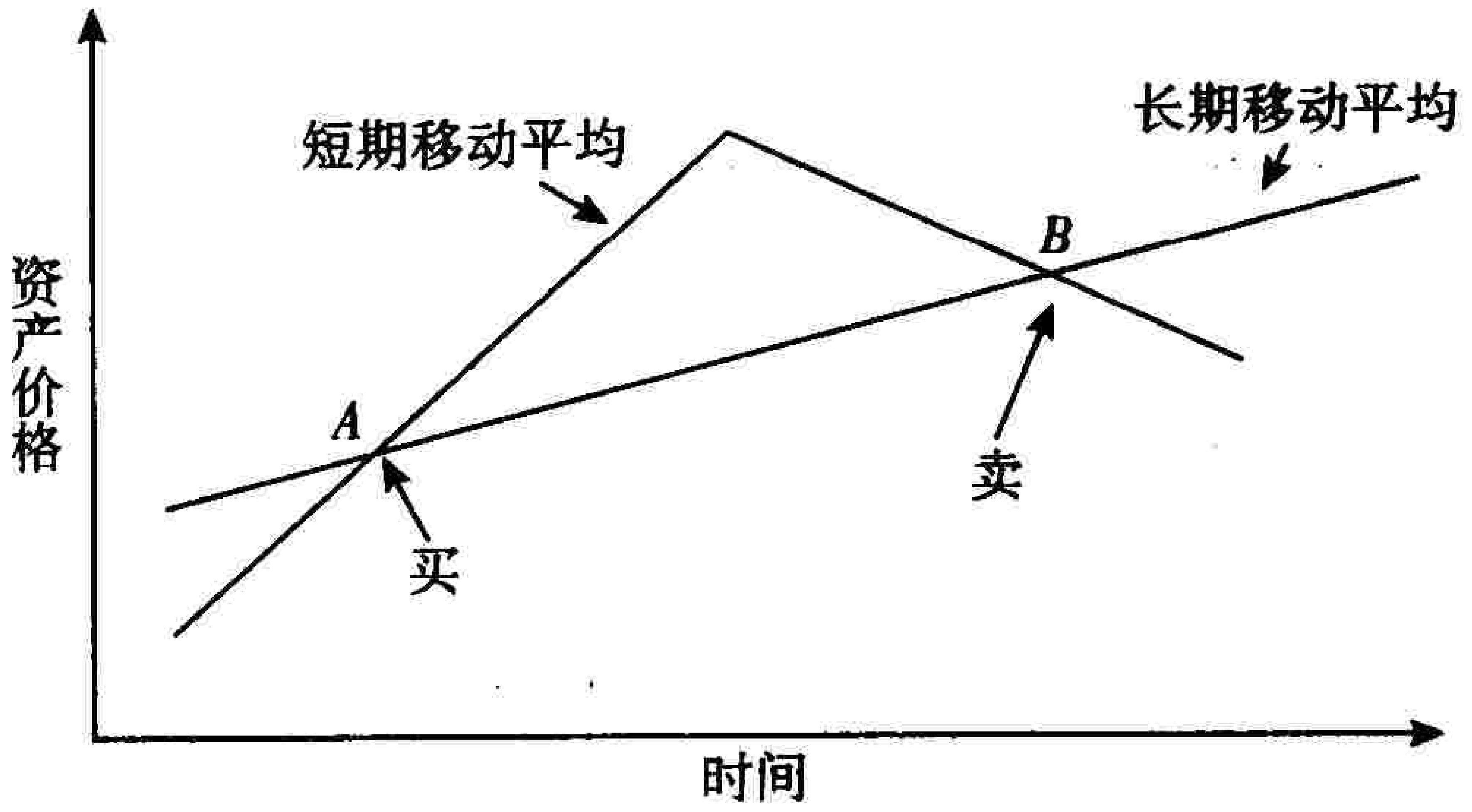


图 3　交叉信号策略

图表专家使用多种方法来预测未来价格变化。例如，他们可能使用历史价格的移动平均值来预测未来价格。他们可能拥有高频图表，比如说每分钟的价格运动图形，试图对图形中的系统模式进行推断。例如，考虑图 1 中的理想模式，这就是众所周知的“头肩反转图形”。图中的水平线称为颈线，当图形运行到 *D* 点（颈线下的峰值）时，图表专家认为这是一个完全反转的信号。

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

他会卖出货币，相信未来价格会下降，打算将来以更低的价格买回货币。另外一个例子是考虑图2中的所谓“对称三角形”图形，振荡幅度越来越小且在 A 点收敛。对某些图表专家而言，这是价格向上的一个信号。在图3中，价格序列的短期移动平均线在 A 点从下上穿长期移动平均线，这是“买入信号”（B 点的情形与之相反）。

显然，这些图形的解释是主观的。作为影响市场的一个群体，绝大多数图表专家必须按照相同方式解释图表，否则他们就会在价格中引入随机噪声而不是趋势。众所周知，图表专家也会使用“市场情绪”的调查数据。例如，如果所报告的“情绪”是对欧洲经济很乐观，那么图表专家就会提前买进欧元，这也会产生“羊群效应”（bandwagon effect）。

预期形成建模

为了解决汇率预测是否不稳定的问题，我们需要考察汇率预测的异质性问题。预期异质性可以用如下式子来进行检验：

$$S^e_{j,t} - S^e_{A,t} = g_j + \varepsilon_{j,t} \tag{42}$$

其中，$S^e_{j,t}$表示个体预期，$S^e_{A,t}$横截面平均预测值。Ito（1990）及 Elliot 和 Ito（1999）发现以日本为基础的汇率预测值存在很强的异质性（也就是说，$g_j \neq 0$），MacDonald 和 Marsh（1996）发现其他货币也有类似结果。

接下来，我们需要考察投资者使用的是那类预测过程，因为这会决定羊群效应是否适用，也会决定预期是否是均值反转的。外推模型给出的羊群效应为（对 $\beta_{j,j} > 0$）：

$$S^e_{j,t,h} - S_t = \alpha_{j,h} + \beta_{j,h}(S_t - S_{t-1}) + \delta_{j,h}Z_t + \varepsilon_{j,t,h} \tag{43}$$

这里的 $S^e_{j,t,h}$是个体 j 在 t 时期的 h 期（如3个月）预期汇率变化。Z_t 包含一些条件变量（如时间趋势、危机虚拟变量），方程中还包含滞后项 $S_{t-k} - S_{t-k-1}$。

回归模型为

$$S^e_{j,t,h} - S_t = \alpha_{j,h} + d_{j,h}(\bar{S}_t - S_{t-1}) + \delta_{j,h}Z_t + \varepsilon_{j,t,h} \tag{44}$$

$0 < d_{j,h} < 1$，$\bar{S}_t$ 是均衡汇率的一个度量指标，它等于汇率的历史平均值或根据基本经济变量（如PPP）来进行计算。

自适应性模型（adaptive model）为：

$$S^e_{j,t,h} - S^e_{j,t-h,h} = \alpha_{j,h} + (1 + f_{j,h})(S_t - S^e_{j,t-h,h}) + \delta_{j,h}Z_t + \varepsilon_{j,t,h} \tag{45}$$

$0 < (1 + f) < 1$。“随机效应”模型允许“斜率”和截距系数在不同个体之间所有变化，而“固定效应”模型要求斜率系数对所有个体都相同。

Bénassy - Quérér、Larribeau 和 MacDonald（2002）使用了三种货币期限为 $h = 3$ 个月和12个月的预测数据，35位参与者在1990年1月至1994年12月期间每周被调查一次。研究发现，绝大多数预测者使用外推预测，但不同预测者对参数 $\beta_{j,h}$的选择不同，尽管对3个月和12个月期限而言，所有值都在 -0.3 和 -0.41 之间。后者意味着汇率上升之后会出现下一期的预期汇率下降，早期研究基本上都发现了这一结论（Cavaglia、Verschoor 和 Wolff（1993a，b），Prat 和 Uctum（1994））。一些预测者确实使用了回归模型和自适应模型，因此所使用的模型也存在差异。除了模型差异和参数估计值不同（对使用相同模型的不同预测者而言）等这些表明证据外，我们不可能对这些异质性的经济意义进行衡量。

当将这三个模型都包含在一个单方程中时，这意味着一些人在两个不同的期限中（如3个月和12个月）使用了不同的模型，没有哪个模型能主导所有预测者的结果（尽管很多人现在使用的是回归模型）。

Allen - Taylor 在研究中所使用的数据是在1988年6月到1989年3月期间对一小群图表专家（每周有10到20个）进行调查后的数据。被调查者每个周四都会接到电话，询问他们对未来一到四周期间的英镑 - 美元、美元 - 德国马克和美元 - 日元汇率的预期值，每一位被调查者对每种货币都有36个观测值。被调查者还会被问到预测时使用了那类信息，将这些信息传递给了谁（如真实的交易者）。

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

研究发现,在最短的期限内,比方说一天内到一个星期,约有90%的被调查者使用图表信息来形成汇率预期。当期限延长到3个月、6个月和1年时,基本经济变量信息的重要性所有增加,从长期来看85%的被调查者认为"基本面分析"比图表分析更为重要。然而,图表分析也可以视为基本面分析的补充,因此,图表分析甚至在较长的期限中对汇率有所影响。

如果事后对图表专家整体的预测精度进行考察,那么一般都会发现Allen和Taylor所发现的如下结果:

* 一种倾向是在预测中漏掉了转折点。在一个上升或下降市场中,图表专家的预期会低估上升或下降幅度。

* 期限为4周的预测误差大于期限为1周的预测误差。个体提前四周的预测一般是无偏的,而提前一周的预测是有偏的。

* 就图表专家整体而言,能正确预测未来一周和四周汇率变化的概率近似为50%,因此,我们认为他们能进行正确预测纯粹是一种偶然。

然而,图表专家整体的上述结果忽略了个体有可能在一段时间上都做得很好的事实。事实上,就预测精度而言,图表专家之间存在差异,一些人确实做得很好。然而,我们从最后的结论中不可能知道得很多,因为调查时期相当短,调查对象也是随机选取的,人们总是期望部分人"比平均水平更好"。

再一次将图表专家视为一个整体,Allen和Taylor就他们的预测绩效是否比其他方法好进行了考察(如基于随机游走模型、ARIMA模型和使用汇率、利率差异、相对股票市场绩效的Var模型等进行预测)。结果不尽一致。然而,部分专家确实能够战胜随机游走模型。在绝大多数情形中,ARIMA和Var预测结果都比随机游走"无变化"模型的预测结果差,绝大多数图表专家无法战胜统计预测模型。但这并不是结论的全部。就提前1周和4周的预测结果来看,统计预测方法和图表专家预测有着相近的均方根误差,总的来说随机游走模型的结果最好。然而,还是有一些图表专家的绩效持续地好于其他模型。

因为Allen和Taylor拥有预期数据,所以他们能考察预期变化与实际汇率变化之间的相关性。我们特别感兴趣的是图表专家是否具有羊群效应式的预期。也就是说,当汇率t-1在t到之间增加时,图表专家是否会向上修正预期? Allen和Taylor对这一假设进行了检验,结果却发现对图表专家整体而言羊群效应式的预期不适用。因此,图表专家的建议从本质来看似乎不会改变,因为他们不会对汇率的近期变化产生过度反应。Allen和Taylor还考察了图表专家是否具有适应性预期或回归预期。从本质来看,这种预期就是均值反转预期,一些图表专家确实具有这种行为。总的来说,研究结果似乎表明一些投资者在市场上会产生系统预测误差,但这种行为似乎没有"羊群效应",图表专家至多对汇率与基础变量之间的偏差有所影响。

Allen-Taylor的研究没有考察图表专家的预测结果是否能产生利润,他们只是对图表专家的预测精度进行了考察。然而,已有大量研究对预测结果的事后价值进行了评估(Goodman(1979,1980),Levich(1980),Bilson(1981)),一些研究是由技术分析师(如图表专家)来进行的。这些研究的主要发现是,一些外汇咨询服务的预测绩效确实比未来即期汇率的远期汇率利率预测指标好。

对其他外汇交易市场参与者的调查结果也表明技术分析在短期内适用,随着期限的增加,基本面分析所起的作用越大(如Lui和Mole(1998)对香港外汇交易经纪商的调查研究、Cheung和Wong(1999,2000)对香港、东京、新加坡外汇交易市场的调查研究,Menkhoff(1998)对德国外汇交易经纪商的调查研究)。

总而言之,对外汇交易者的调查结果表明预期存在异质性,但几乎没有发现很强的羊群效应,羊群效应可以使即期汇率在短期内产生剧烈波动。这一证据使得即期汇率短期变化的解释变得更加复杂(例如,这是风险效应、比索效应还是无法解释的非线性反应?)。

调查对考察技术交易在各种市场上是否适用是非常有用的,定量和定性预期调查数据可以用来对理性预期假设的适用性进行评估。但是,这些证据没有说明这些因子在价格决定中到底

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

有多重要，但为假设不同类型交易者、学习机制和有限理性形式的其他模型提供了支持。

29.6　技术交易法则

过滤器法则也可以用来预测汇率。一种简单的过滤器法则是在汇率相对于近期底部上涨了后买进货币，在汇率相对于近期高峰值下降时卖出货币。如果市场有效且 UIP 成立，那么执行这一策略的利息成本应等于即期汇率变化所产生的利润。早期研究（如 Dooley 和 Shafer（1983），Levich 和 Thomas（1993））发现，对几日或几周使用过滤器法则确实能获得利润，Engle 和 Hamilton（1990）的近期研究指出，美元在 20 世纪 70 年代和 80 年代具有长期趋势波动特征，使用“趋势跟踪”过滤期法则可以获得利润。当然，在各个子时期中也可能遭致损失，利润并不是没有风险。因此，使用每日数据和每周数据发现的结果是，即使对交易成本进行了考察，技术分析法则也可以在外汇市场上获利，我们很难明白风险是如何进行补偿的（如 Levich 和 Thomas（1993），Neely、Weller 和 Dittmar（1997））。然而，这些研究中的交易频率大概是每年 3 到 26 次。但调查证据（Allen 和 Laylor（1990），Cheung 和 Chinn（2001））却表明技术分析在非常短的时期中适用（也就是说，几天内的交易或日内交易（技术交易的目的是在当日结束时持有零的净头寸））。因此，在非常短的时期中赚取利润（对风险进行调整）是否可能呢？

Goodhart 和 Curcio（1992）使用的是路透社发布的分析师预测支撑价位和阻力价位，Curcio、Goodhart、Guillaume 和 Payne（1997）使用日内数据来对过滤器法则进行检验，但没有发现利润机会。苏黎世 Olsen & Associates 认为私人（未公开的）交易模型使用五分钟数据可以获得利润。

Neely 和 Weller（2003）在最近的研究中使用的是四种货币（英镑、马克、日元和瑞士法郎）与美元之间即期汇率每半小时的数据（目的是为了避免一些微观结构问题，如买卖反弹等，Lyons（2001））。预测模型基于如下数据使用遗传算法：（ⅰ）$S_T\sqrt{S_t}$，$\bar{S}$ 是前两周的移动平均值；（ⅱ）近期期货价格差异（美国的期货价格减去其他外币期货价格）；（ⅲ）天数。人们发现后两者其实没有提供什么信息。t 时刻的信号 z_t 在 $[0,T]$ 上的累积超额收益为：

$$r = \sum_{t=0}^{T-1} z_t r_t + n\ln(1-2c)$$

其中，$z_t=+1(-1)$ 是 t 时刻的多头（空头），r_t 是 $[t,t+1]$ 期间持有国外货币多头的连续复利收益，n 是 $[0,T]$ 期间的交易量。常数 c 表示单向交易成本（在 1 到 5 个基点之间）。针对两个数据时期（以 1996 年 5 月 31 日为结束日）对遗传算法进行估计（“训练”）和“筛选”，由此来产生最大的 r 值。样本外预测期为 1996 年 6 月 1 日 ~1996 年 12 月 31 日。

Neely 和 Weller 所使用的替代模型是线性自回归过程和一个过滤器法则。例如，如果你在 $t-1$ 时刻持有一个多头头寸，那么当模型所预测的下降幅度大于过滤值 $f(=1,2,\cdots,$或 5 个基点）时，你会在 $t-1$ 时刻转而持有一个空头头寸。如果遗传算法的训练是在零交易成本假设下进行的，那么“样本外”预测会产生一个高达 100% 的年度收益。但这涉及每个小时的交易，单边交易存在 1 个基点左右的交易成本（对大额外汇交易而言实际成本约为 2 到 2.5 个基点），因此，遗传算法预测在考虑交易成本后无法获得利润。如果是在 $c=1$ 或 2 个基点时来对遗传算法进行训练，那么样本外期间的交易频率会急剧下降，在扣除交易成本后交易仍然无利可图。使用自回归过滤器法则也会发现类似结果，部分原因是遗传算法主要取决于汇率变化的滞后值（而不是另外两个变量），因此它与自回归模型较为相似。上述证据表明短期外汇交易“没有免费的午餐”。

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

数据探测偏差

Sullivan、Timmermann 和 White(1999)注意到许多研究都发现技术交易法则可以赚取利润,他们强调的是这是否是一个"偶然性"(数据挖掘)问题。文献中报告的结果都是比较"成功"的结果,因此这些结果是有偏的。基本思想是考虑如下问题,为了可预测性而搜寻最大的统计量值,以此为基础的检验规模与正常值之间可能存在很大的差异。他们使用 White(2000)的"自助探测法",一旦数据挖掘效应被证实,那么这种方法能给出特殊交易法则"成功"的 p 值。宽泛地说,数据探测思想试图回避幸存偏差问题(也就是说,我们在学术期刊上可以看到可预测性研究方面的文章不成比例)。

Sullivan、Timmermann 和 White(1999)使用道琼斯工业平均指数(DJIA)在过去 100 年中的每日数据和标准普尔 500 期货合约在 1984—1996 年期间每日数据对技术交易法则(如过滤器法则、移动平均法、支撑价位和阻力价位等)进行了考察。他们考虑了这些交易策略的 8000 多个变体。有证据表明,在 1976—1986 年期间,即使对数据探测问题进行了调整,一些交易法则的绩效也比基准投资组合(即持有现金)好。但是,绩效最好的交易法则在 1987—1996 年的样本外时期中表现优于基准投资组合的概率仅仅为 12%,根据标准临界值判断的结果是这种绩效在统计上不显著。因此,在过去 10 年中市场似乎越来越有效,可能的原因是信息加工更加便宜也更加快速。(我们在异像一章中已经对日历法则的数据探测偏差问题进行了讨论。)

考虑数据探测偏差的思想导致了如下结果,对标准普尔 500 期货指数而言,特殊交易法则在 1984—1996 年期间能获得每年 10% 收益优良绩效(相对于现金而言)的 p 值为 0.04。但是如果从大量可能交易法则中选定某个交易法则,那么数据探测的 p 值实际上为 0.90。这说明数据挖掘可能有很大的影响(如果你认为你发现了一种能赚取百万美元且在统计上显著的交易法则,那么你应该看看 Sullivan 等人的这篇文章)。

29.7 小结

* 在过去 10 年中,使用 ARCH 和 GARCH 过程来考察资产收益条件波动率(条件方差)的实证研究出现了快速增长。使用美国数据所得到的实证结果是,在股票收益条件波动率中只有很小一个部分可以用历史波动率和基本经济变量(如产出等)的波动率来进行解释。

* 大量证据表明股票和外汇(日内、每日、每周和每月)收益的条件方差存在持续性。在更长的收益期中人们通常不会发现 GARCH 效应。

* 股票收益条件波动率的强持续性意味着所观察到的股票价格运动可能与具有时变风险溢价的 RVF 相一致。然而,条件方差对预期市场收益的影响很难用 GARCH - M 模型来捕获。(市场投资组合某个子集的)条件方差对收益的影响与之类似。

* 可能出现的情形是,噪声交易者和精明交易一起影响股票市场综合指数的预期收益,噪声交易者的影响取决于我们是处于"高"波动时期还是"低"波动时期。

* 就短期零息债券而言,使用 GARCH - M 模型所发现的结果是,时变期限溢价影响 HPR 的证据非常微弱。国库券的价格变化在绝大多数时期中都比长期债券价格的变化小,因此,上述结果并不令人感到奇怪。在短期国债市场经历了剧烈波动后(如美国 1979—1982 年),时变期限溢价的持续性证据更为明显,条件方差对预期收益的影响也更为明显。

* 长期债券的持有期收益似乎要受时变条件方差和协方差的影响,但这种关系是否稳定仍然是一个公开问题。

* 尽管 ARCH 和 GARCH 模型为时变二阶矩的建模提供了有用的统计基础,一些法则的参数化技巧(特别是在多变量模型中)使得实证研究经常无法进行精确估计。

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

- 市场微观结构研究，特别是对外汇交易市场的研究，已经说明了指令流作为即期汇率变化近似决定因素的重要性，也说明了存货持有量作为买卖价差关键决定因素的重要性。
- 噪声交易者（图表专家）确实会对短期中的即期外汇汇率产生影响，但调查证据表明图表专家和其他外汇交易预测者不太可能形成使汇率出现更大变动的预期。然而，因为调查数据无法度量预测者的“市场影响”（即他们的预测能产生多大的“货币”回报），所以任何推断至多是定性推断。
- 技术交易法则应用于股票和即期外汇市场的高频数据（如日内收益、每日收益）可以产生超额收益（扣除交易成本并对风险进行调整）的证据似乎不是特别的强。你必须辛苦地工作才能找到潜在的获利“策略”，重复博弈的交易成本和风险都会很高。

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

参考文献

1. Abel, A. B. (1990) Asset Prices under Habit Formation and Catching up with the Joneses, American Economic Review, 80, 38 - 42.

2. Admati, A. and Pfleiderer, P. (1988) A Theory of Intraday Patterns: Volume and Price Variability, Review of Financial Studies, 1, 3 - 40.

3. Agarwal, V. and Naik, N. Y. (2000) Multi - period Performance Persistence Analysis of Hedge Funds, Journal of Financial and Quantitative Analysis, 35, 327 - 342.

4. Ali, A., Hwang L. - S. and Trombley, M. A. (2003) Arbitrage Risk and the Book - to - market Anomaly, Journal of Financial Economics, 69, 355 - 373.

5. Allen, D. and Tan, M. (1999) A Test of the Persistence in the Performance of UK Managed Funds, Journal of Business Finance and Accounting, 26(5), 559 - 593.

6. Allen, H. and Taylor, M. P. (1989) Chart Analysis and the Foreign Exchange Market, Bank of England Quarterly Bulletin, 29(4), 548 - 551.

7. Allen, H. and Taylor, M. P. (1990) Charts, Noise and Fundamentals in the Foreign Exchange Market, Economic Journal, 100(400), 49 - 59.

8. Anderson, E., Hansen, L. and Sargent, T. (2000) Robustness, Detection and the Price of Risk, Working Paper, NYU.

9. Anderson, N., Breedon, F. and Deacon, M. (1996) Estimating and Interpreting the Yild Curve, John Wiley, Chichester.

10. Anderson, N. and Sleath, J. (1999) New Estimates of the UK Nominal and Real Yield Curves, Bank of England Quarterly Bulletin, 39(4), 384 - 392.

11. Andreassen, P. B. and Kraus, S. J. (1990) Judgmental Extrapolation and the Salience of Change, Journal of Forecasting, 9(4), 347 - 372.

12. Ang, A. and Bekaert, G. (2001) Stock Return Predictability - Is it There? Working Paper, Columbia University, New York.

13. Ang, A. and Piazzesi, M. (2003) A No - arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables, Journal of Monetary Economics, 50(4), 745 - 787.

14. Ardeni, P. G. and Lubian, D. (1991) Is There Trend Reversion in Purchasing Power Parity, European Economic Review, 35(5), 1035 - 1055.

15. Arrow, K. J (1970) Essays in the Theory of Risk Bearing, North Holland, Amsterdam, The Netherlands.

16. Asch, S. E. (1952) Social Psychology, Prentice Hall.

17. Attanasio, O. and Wadhwani, S. (1990) Does the CAPM Explain Why the Dividend Yield Helps Predict Returns?, London School of Economics, Financial Markets Group Discussion Paper No. 04.

18. Backus, D., Foresi, S. and Telmer, C. (2001) Affine Term Structure Models and the Forward Premium Anomaly, Journal of Finance, 56, 279 - 304.

19. Backus, D., Foresi, S., Mozumdar, A. and Wu, L. (2001) Predictable Changes in Yields and Forward Rates, Journal of Financial Economics, 59, 281 - 311.

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

20. Backus, D., Gregory, A. and Telmer, C. (1993) Accounting for Forward Rates in Markets for Foreign Currency, Journal of Finance, 48(5), 1887 - 1908.

21. Baillie, R. T. and Bollerslev, T. (1989) Common Stochastic Trends in a System of Exchange Rates, Journal of Finance, 44(1), 167 - 181.

22. Baillie, R. T. and McMahon, P. C. (1989) The Foreign Exchange Market: Theory and Evidence, Cambridge University Press, Cambridge, Mass.

23. Baker, M. and Wurgler, J. (2000) The Equity Share in New Issues and Aggregate Stock Returns, Journal of Finance, 55(5), 2219 - 2257.

24. Baker, M. and Wurgler, J. (2002) Market Timing and Capital Structure, Journal of Finance, LVII, 1 - 32.

25. Bakshi, G. S. and Chen, Z. (1996) The Spirit of Capitalism and Stock Market Prices, American Economic Review, 86(1), 133 - 157.

26. Balfoussia, C. and Wickens, M. (2003) Macroeconomic Sources of Risk in the Term Structure, mimeo, Department of Economics, University of York, February.

27. Balke, N. S. and Fomby, T. B. (1997) Threshold Cointegration, International Economic Review, 38, 627 - 645.

28. Balke, N. S. and Wohar, M. E. (1998) Nonlinear Dynamics and Covered Interest Parity, Empirical Economics, 23, 535 - 559.

29. Banerjee, A., Dolado, J. J., Galbraith, J. W. and Hendry, D. F. (1993) Co - integration, Error - correction, and the Econometric Analysis of Non - stationary Data, Oxford University Press.

30. Bank for International Settlements. (2001) Central Bank Survey of Foreign Exchange and Derivatives Market Activity, Basle.

31. Banz, R. W. (1981) The Relationship Between Return and Market Value of Common Stock, Journal of Financial Economics, 9(1), 3 - 18.

32. Barber, B. and Lyon, J. (1997) Detecting Long - run Abnormal Stock Returns: The Empirical Power and Specification of Test Statistics, Journal of Financial Economics, 43, 341 - 372.

33. Barber, B. M. and Odean, T. (2000) Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors, Journal of Finance, 55, 773 - 806.

34. Barberis, N. (2000) Investing for the Long Run when Returns are Predictable, Journal of Finance, 55, 225 - 264.

35. Barberis, N. and Huang, M. (2001) Mental Accounting, Loss Aversion, and Individual Stock Returns, Journal of Finance, 56, 1247 - 1292.

36. Barberis, N. and Shleifer, A. (2003) Style Investing, Journal of Financial Economics, 68, 161 - 199.

37. Barberis, N. and Thaler, R. (2003) A Survey of Behavioural Finance, in G. M. Constantinides, M. Harris and R. Stulz (eds.), Handbook of the Economics of Finance, Elsevier Science B. V.

38. Barberis, N., Huang, M. and Santos, T. (2001) Prospect Theory and Asset Pricing, Quarterly Journal of Economics, CX ~ I(1), 1 - 53.

39. Barberis, N., Shleifer, A. and Wurgler, J. (2001) Comovement, Unpublished Working Paper, University of Chicago.

40. Barberis, N., Shleifer, A. and Vishny, R. (1998) A Model of Investor Sentiment, Journal of Financial Economics, 49, 307 - 345.

41. Barclays Capital. (2003) Equity Gilt Study, Barclays Capital, London.

42. Barkoulas, J., Baum, C. F. and Chakraborty, A. (2003) Forward Premiums and Market Efficiency: Panel Unit - root Evidence from the Term Structure of Forward Premiums, Journal of Macroeconomics, 25, 109 - 122.

43. Barsky, R. B. and DeLong, J. B. (1993) Why Does the Stock Market Fluctuate?, Quarterly Journal of Economics, 108(2), 291 - 311.

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

44. Bartholdy, J. and Peare, P. (2003) Unbiased Estimation of Expected Return Using CAPM, International Review of Financial Analysis, 12, 69 - 81.

45. Batchelor, R. A. and Dua, P. (1987) The Accuracy and Rationality of UK Inflation Expectations: Some Qualitative Evidence, Applied Economics, 19(6), 819.

46. Becker, G. S. (1991) A Note on Restaurant Pricing and Other Examples of Social Influences on Price, Journal of Political Economy, 99(5), 1109 - 1116.

47. Bekaert, G. and Hodrick, R. J. (1993) On Biases in the Measurement of FX Risk Premiums, Journal of International Money and Finance, 12, 115 - 138.

48. Bekaert, G. (1996) The Time - variation of Risk and Return in Foreign Exchange Markets: A General Equilibrium Perspective, Review of Financial Studies, 9, 427 - 470.

49. Bekaert, G. and Hodrick, R. J. (1992) Characterising Predictable Components in Excess Returns on Equity and Foreign Exchange Markets, Journal of Finance, 47(2), 467 - 509.

50. Bekaert, G. and Hodrick, R. J. (2001) Expectations Hypothesis Tests, Journal of Finance, 56, 1357 - 1394.

51. Bekaert, G., Hodrick, R. J. and Marshall, D. A. (1997) On Biases in Tests of the Expectations Hypothesis of the Term Structure of Interest Rates, Journal of Financial Economics, 44(3), 309 - 348.

52. Bekaert, G., Hodrick, R. J. and Marshall, D. A. (2001) Peso Problem Explanations for Term Structure Anomalies, Journal of Monetary Economics, 48(2), 241 - 270.

53. Benartzi, S. (2001) Excessive Extrapolation and the Allocation of 401(k) Accounts to Company Stock, Journal of Finance, 56(5), 1747 - 1764.

54. Benartzi, S. and Thaler, R. H. (1995) Myopic Loss Aversion and the Equity Premium Puzzle, Quarterly Journal of Economics, CX(1), 73 - 92.

55. Benartzi, S. and Thaler, R. H. (2001) Naive Diversification Strategies in Defined Contribution Saving Plans, American Economic Review, 91(1), 79 - 98.

56. Benassy - Querer, A., Larribeau, S. and MacDonald, R. (2002) Models of Exchange Rate Expectations: How Much Heterogeneity? International Financial Markets, Institutions and Money, 13, 113 - 136.

57. Bergin, P. R. (2003) Putting the New Open Macroeconomics to a Test, Journal of International Economics, 60, 3 - 34.

58. Bergin, P. R. and Feenstra, R. C. (2001) Pricing - to - market, Staggered Contracts and Real Exchange Rate Persistence, Journal of International Economics, 54, 333 - 359.

59. Berk, J. (1997) Does Size Really Matter? Financial Analyst Journal, Sept. /Oct., 12 - 18.

60. Berkowitz, S. A., Logue, D. E. and Noser, E. A. (1988) The Total Costs of Trading in the NYSE, Journalof Finance, 43, 97 - 112.

61. Bernard, V. L. and Thomas, J. (1989) Post - earnings Announcement Drift: Delayed Price Response or Risk Premium?, Journal of Accounting Research, 27, 1 - 48.

62. Bernard, V. L. and Thomas, J. (1990) Evidence that Stock Prices do not Fully Reflect the Implications of Current Earnings for Future Earnings, Journal of Accounting and Economics, 13, 305 - 340.

63. Bernstein, R. (1995) Style Investing, John Wiley, New York.

64. Bessembinder, H. (1994) Bid - ask Spreads in the Interbank Foreign Exchange Markets, Journal of Financial Economics, 35, 317 - 348.

65. Betts, C. and Devereux, M. B. (2000) Exchange Rate Dynamics in a Model of Pricing - to - Market, Journal of International Economics, 50(1), 215 - 244.

66. Bhamra, H. and Uppal, R. (2002) Non - Redundant Derivatives in a Dynamic General Equilibrium Economy, mimeo, London Business School.

67. Bilson, J. F. O. (1978) The Monetary Approach to the Exchange Rate: Some Empirical Evidence, IMF

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

Staff Papers,25,48 – 77.

68. Bilson,J. F. O. (1981) The "Speculative Efficiency" Hypothesis,Journal of Business,54,435 – 451.

69. Black,F. (1972) Capital Market Equilibrium with Restricted Borrowing,Journal of Business,45(3),444 – 455.

70. Black,F. and Scholes,M. (1973) The Pricing of Options and Corporate Liabilities,Journal of Political Economy,81(3),637 – 654.

71. Black,F.,Jensen,M. C. and Scholes,M. (1972) The Capital Asset Pricing Model: Some Empirical Tests, in M. C. Jensen (ed.),Studies in the Theory of Capital Markets,Praeger,New York.

72. Blake,D. and Timmermann,A. (1998) Mutual Fund Performance: Evidence from the UK, European Finance Review,2,57 – 77.

73. Blake,D.,Lehman,B. and Timmermann,A. (1999) Asset Allocation Dynamics and Pension Fund Performance,Journal of Business,72,429 – 461.

74. Blanchard,O. J. (1979) Speculative Bubbles,Crashes and Rational Expectations,Economic Letters,3,387 – 389.

75. Blanchard,O. J. and Watson,M. W. (1982) Bubbles, Rational Expectations and Financial Markets, in P. Wachtel (ed.), Crises in the Economic and Financial System, Lexington Books, Lexington, Mass., pp. 295 – 315.

76. Blume,M. (1975) Betas and their Regression Tendencies,Journal of Finance,30(3),785 – 795.

77. Bodie,Z.,Merton,R. C. and Samuelson P. (1992) Labour Supply Flexibility and Portfolio Choice in a Life – cycle Model,Journal of Economic Dynamics and Control,6(3/4),725.

78. Bodurtha,J.,Kim,D. and Lee,C. M. (1993) Closed – end Country Funds and US Market Sentiment, Review of Financial Studies,8,879 – 918.

79. Bollerslev,T.,Chou,R. Y. and Kroner,K. F. (1992) ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence,Journal of Econometrics,52,5 – 59.

80. Bollerslev,T.,Engle,R. F. and Wooldridge,J. M. (1988) A Capital Asset Pricing Model with Time – varying Covariances,Journal of Political Economy,96(1),116 – 131.

81. Boudoukh,J.,Richardson,M. and Whitelaw,R. (1995) Expect the Worst,Risk,8,100 – 101.

82. Brandt,M. W. (1998) Estimating Portfolio Consumption Choice: A Conditional Euler Equations Approach?,Wharton School W. P.

83. Branson,W. H. (1977) Asset Markets and Relative Prices in Exchange Rate Determination, Sozial Wissenschaftliche Annalen,1.

84. Brav,A. (2000) Inference in Long – Horizon Event Studies,Journal of Finance,55,1979 – 2016.

85. Breeden,D. T.,Gibbons,M. R. and Litzenberger,R. H. (1989) Empirical Tests of the Consumption – oriented CAPM,Journal of Finance,44(2),231 – 262.

86. Bremer,M. A. and Sweeney,R. J. (1991) The Reversal of Large Stock Price Decreases, Journal of Finance,46(2),747 – 754.

87. Brennan,M. J.,Schwartz,E. S. and Lagnado,R. (1997) Strategic Asset Allocation,Journal of Economic Dynamics and Control,21,1377 – 1403.

88. Breuer,J. B. and Wohar,M. E. (1996) The Road Less Travelled: Institutional Aspects of Data and their Influence on Empirical Estimates with an Application to Tests of Forward Rate Unbiasedness,Economic Journal, 106(434),26 – 38.

89. Britten – Jones,M. (1999) The Sampling Error in Estimates of Mean – variance Efficient Portfolio Weights,Journal of Finance,52(2),637 – 659.

90. Brooks,C. (1998) Forecasting Stock Return Volatility: Does Volume Help?,Journal of Forecasting,17,59 – 80.

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

91. Brooks, C. and Persand, G. (2001) The Trading Profitability of Forecasts of the Gilt – equity Yield Rates, International Journal of Forecasting, 17, 11 – 29.

92. Brooks, C., Henry, O. T. and Persand, G. (2002) Optimal Hedging and the Value of News, Journal of Business, 75(2), 3330 – 3352.

93. Brown, G., Draper, P. and McKenzie, E. (1997) Consistency of UK Pension Fund Investment Performance, Journal of Business Finance and Accounting, 24(2), 155 – 178.

94. Brown, S. J., Goetzmann, W. N. and Ibbotson, R. G. (1999) Offshore Hedge Funds Survival and Performance 1989 – 1995, Journal of Business, 72, 91 – 118.

95. Brown, S. J., Goetzmann, W. N. and Park, J. (2001) Careers and Survival Competition and Risk in the Hedge Fund and CTA Industry, Journal of Finance, 56, 1869 – 1886.

96. Brunnermeier, M. K. (2001) Asset Pricing Under Asymmetric Information: Bubbles, Crashes, Technical Analysis and Herding, Oxford University Press, Oxford.

97. Buhlmann, P. (1997) Sure Bootstrap for Time Series, Benoulli, 3, 123 – 148.

98. Bulkley, G. and Taylor, N. (1996) A Cross – section Test of the Present Value Model, Journal of Empirical Finance, 2(4), 295 – 306.

99. Burnside, C. (1994) Hansen – Jagannathan Bounds as Classical Tests of Asset Pricing Models, Journal of Business and Economic Statistics, 12, 57 – 79.

100. Buse, A. (1982) The Likelihood Ratio, Wald and Lagrange Multiplier Test: An Expository Note, American Statistician, 36(3), 153 – 157.

101. Butler, J. S. and Schachter, B. (1996) Improving Value – at – Risk Estimates by Combining Kernel Estimation with Historical Simulation, Proceedings of the 32nd Conference on Bank Structure and Competition, Federal Reserve Bank of Chicago, Chicago, pp. 363 – 380.

102. Camerer, C. and Weber, M. (1992) Recent Developments in Modelling Preferences: Uncertainty and Ambiguity, Journal of Risk and Uncertainty, 5, 325 – 370.

103. Campbell, J. and Clarida, R. H. (1987) The Term Structure of Euromarket Rates: An Empirical Investigation, Journal of Monetary Economics, 19, 25 – 44.

104. Campbell, J. Y. (1986) A Defense of Traditional Hypothesis About the Term Structure of Interest Rates, Journal of Finance, 41, 183 – 193.

105. Campbell, J. Y. (1991) A Variance Decomposition for Stock Returns, Economic Journal, 101(405), 157 – 179.

106. Campbell, J. Y. (1993) Intertemporal Asset Pricing Without Consumption Data, American Economic Review, 83, 487 – 512.

107. Campbell, J. Y. (2000) Asset Pricing at the Millennium, Journal of Finance, 55, 1515 – 1567.

108. Campbell, J. Y. (2001) Forecasting US Equity Returns in the 21st Century, Harvard University, mimeo.

109. Campbell, J. Y. and Cochrane, J. H. (2000) Explaining the Poor Performance of Consumption – based Asset Pricing Models, Journal of Finance, 55, 2863 – 2878.

110. Campbell, J. Y. and Shiller, R. J. (1988) Stock Prices, Earnings, and Expected Dividends, Journal of Finance, 43(3), 661 – 676.

111. Campbell, J. Y. and Shiller, R. J. (1989) The Dividend Ratio Model and Small Sample Bias – A Monte Carlo Study, Economic Letters, 29, 325 – 331.

112. Campbell, J. Y. and Shiller, R. J. (1991) Yield Spreads and Interest Rate Movements: A Birds Eye Review of Economic Studies, 58, 495 – 514.

113. Campbell, J. Y. and Viceira, L. M. (1999) Consumption and Portfolio Decisions When Expected Returns are Time Varying, Quarterly Journal of Economics, 114(2), 433 – 496.

114. Campbell, J. Y. and Viceira, L. M. (2000) Consumption and Portfolio Decisions When Expected Returns

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

are Time Varying: Erratum, Unpublished Paper, Harvard University.

115. Campbell, J. Y., Chan, Y. L. and Viceira, L. M. (2003) A Multivariate Model of Strategic Asset Allocation, Journal of Financial Economics, 67(1), 41 – 80.

116. Campbell, J. Y., Lo, A. W. and MacKinlay, A. C. (1997) The Econometrics of Financial Markets, Princeton University Press.

117. Capocci, D. and Hubner, G. (2004) Analysis of Hedge Fund Performance, Journal of Empirical Finance, 11, 55 – 89.

118. Carhart, M. (1995) Survivor Bias and Mutul Fund Performance, Working Paper, School of Business Administration, University of Southern California, Los Angeles, Calif.

119. Carhart, M. (1997) On Persistence in Mutual Fund Performance, Journal of Finance, 52(1), 57 – 82.

120. Cavaglia S., Melas, D. and Miyashiuta, O. (1994) Efficiency Across Frontiers, Risk, 7(10), 56 – 61.

121. Cavaglia, S., Verschoor, W. F. C. and Wolff, C. P. (1993a) Further Evidence on Exchange Rate Expectations, Journal of International Money and Finance, 12, 78 – 98.

122. Cavaglia, S., Verschoor, W. F. C. and Wolff, C. P. (1993b) On the Biasedness of Forward Foreign Exchange Rates: Irrationality or Risk Premium, Journal of Business, t17, 321 – 343.

123. Cecchetti, S. G., Lam, P. – S. and Mark, N. C. (1990) Mean Reversion in Equilibrium Asset Prices, American Economic Review, 80(3), 398 – 418.

124. Cerny, A. (2004) Mathematical Techniques in Finance: Tools for Incomplete Markets, Princeton University Press.

125. Chacko, G. and Viceira, L. M. (1999) Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets, NBER Working Paper No. 7377.

126. Chalmers, J., Edelen, R. and Kadlec, G. (1999) Transaction Cost Expenditures and the Relative Performance of Mutual Funds, Working Paper, University of Oregon.

127. Chan, L. K. C., Jegadeesh, N. and Lakonishok, J. (1996) Momentum Strategies, Journal of Finance, 51, 1681 – 1713.

128. Chan, L. K. C., Karceski, J. and Lakonishok, J. (2000) New Paradigm or Same Old Hype in Equity Investing?, Financial Analysis Journal, 56, 23 – 26.

129. Chari, V. V., Kehoe, P. J. and McGrattan, E. R. (2001) Can Sticky Price Models Generate Volatile and Persistent Real Exchange Rates?, Federal Reserve Bank of Minneapolis Staff Report, No. 277.

130. Chen, H. L., Jegadeesh, N. and Wermers, R. (2000) The Value of Active Mutual Fund Management: An Examination of the Stockholdings and Trades of Fund Managers, Journal Financial Quantitative Analysis, 35(3), 343 – 368.

131. Chen, N., Roll, R. and Ross, S. A. (1986) Economic Forces and the Stock Market, Journal of Business, 59, 383 – 403.

132. Chen, Y. (2003) Discriminating Between Competing STAR Models, Economics Letters, 79, 161 – 167.

133. Cheung, Y. W. and Chinn, M. D. (2001) Currency Traders and Exchange Rate Dynamics: A Survey of the US Market, Journal of International Money and Finance, 20(4), 439 – 471.

134. Cheung, Y. W. and Lai, K. S. (1994) Mean Reversion in Real Exchange Rates, Economics Letters, 46 (3), 251 – 256.

135. Cheung, Y. W. and Wong, C. Y. P. (1999) Foreign Exchange Traders in Hong Kong, Tokyo and Singapore: A Survey Study, Advances in Pacific Basin Financial Markets, 5, 111 – 134.

136. Cheung, Y. W. and Wong, C. Y. P. (2000) Survey of Market Practitioners' View on Exchange Rate Dynamics, Journal of International Economics, 51(2), 401 – 419.

137. Chevalier, J. and Ellison, G. (1999) Are Some Mutual Fund Managers Better Than Others? Cross Sectional Patterns in Behaviour and Performance, Journal of Finance, 54(3), 875 – 899.

www.55188.com最好的股票论坛理想在线证券网欢迎您!

www.55188.com最好的股票论坛理想在线证券网欢迎您!

138. Chew, S. (1989) Axiomatic Utility Theories with Betweeness Property, Annals of Operations Research, 19, 273 – 298.

139. Chiang, R., Davidson, I. and Okunev, J. (1997) Some Further Theoretical & Empirical Implications Regarding the Relationship Between Earnings Dividends & Stock Prices, Journal of Banking and Finance, 21, 17 – 35.

140. Chopra, N., Lakenishok, J. and Ritter, J. (1992) Measuring Abnormal Performance: Do Stocks Overreact? Journal of Financial Economics, 31, 235 – 268.

141. Chou, R. Y. (1988) Volatility Persistence and Stock Valuations: Some Empirical Evidence Using GARCH, Journal of Applied Econometrics, 3, 279 – 294.

142. Christiansen, C. (2003) Testing the Expectations Hypothesis Using Long – maturity Forward Rates, Economics Letters, 78, 175 – 180.

143. Christopherson, J., Ferson, W. and Glassman, D. (1998) Conditioning Manager Alphas on Economic Information: Another Look at the Persistence of Performance, Review of Financial Studies, 11(1), 111 – 142.

144. Clare, A., O'Brien, R., Thomas, S. and Wickens, M. (1993) MacrOeconomic Shocks and the Domestic CAPM: Evidence from the UK Stock Market, Discussion Paper 93 – 02, Department of Economics, Brunel University, UK.

145. Clare, A., Priestly, R. and Thomas, S. (1997) Stock Return Predictability or Mismeasured Risk?, Applied Financial Economics, 7(6), 679 – 688.

146. Clarida, R. and Taylor, M. P. (1997) The Term Structure of Forward Exchange Premiums and the Forecastability of Spot Exchange Rates: Correcting the Errors, Review of Economics and Statistics, 79, 353 – 361.

147. Clarida, R., Gali, J. and Gertler, M. (1999) The Science of Monetary Policy: A New Keynesian Perspective, Journal of Economic Literature, XXXVII(4), 661 – 1734.

148. Clewlow, L. and Strickland, C. (1998) Implementing Derivatives Models, John Wiley, Chichester.

149. Clinton, K. (1988) Transactions Costs and Covered Interest Arbitrage: Theory and Evidence, Journal of Political Economy, 96, 358 – 370.

150. Cochrane, J. H. (1996) A Cross – sectional Test of an Investment – based Asset Pricing Model, Journal of Political Economy, 104, 572 – 621.

151. Cochrane, J. H. (1997) Where is the market Going?, Economic Perspectives, (Federal Reserve Bank of Chicago) 21(6).

152. Cochrane, J. H. (2001) Asset Pricing, Princeton University Press, Princeton, N. J.

153. Cochrane, J. H. and Hansen, L. P. (1997) Asset Pricing Explorations for Macroeconomics in O. Blanchard and S. Fisher (eds), NBER Macroeconomics Annual 1997, MIT Press, Cambridge, Mass., pp. 115 – 165.

154. Constantinides, G. M. (1982) Intertemporal Asset Pricing with Heterogeneous Consumers and without Demand Aggregation, Journal of Business, 55(2), 253 – 267.

155. Constantinides, G. M. (1990) Habit Formation – A Resolution of the Equity Premium Puzzle, Journal of Political Economy, 98(3), 519 – 543.

156. Constantinides, G. M. and Duffle, D. (1996) Asset Pricing with Heterogeneous Consumers, Journal of Political Economy, 1D4(2), 219 – 240.

157. Corradi, V., Swanson, N. R. and White, H. (2000) Testing for Stationarity – ergodicity and for Co – movements Between Non – linear Discrete Markov Processes, Journal of Econometrics, 96(1), 1.

158. Cox, J. C., Ingersoll Jr, J. E. and Ross, S. A. (1981) A Re – examination of Traditional Hypotheses about the Term Structure of Interest Rates, Journal of Finance, 36(4), 769 – 799.

159. Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985) A Theory of the Term structure of Interest Rates, Econometrica, 53, 385 – 407.

160. Cumby, R. (1990) Consumption Risk and International Equity Returns: Some Empirical Evidence, Jour-

www.55188.com最好的股票论坛理想在线证券网欢迎您!

nal of International Money and Finance, 9, 182 - 192.

161. Cumby, R. E. (1988) Is it Risk? Explaining Deviations from Uncovered Interest Parity, Journal of Monetary Economics, 22, 279 - 299.

162. Cumby, R. E. and Huizinga, J. (1992) Investing the Correlation of Unobserved Expectations: Expected Returns on Equity and Foreign Exchange Markets and Other Examples, Journal of Monetary Eco - nomics, 30, 217 - 253.

163. Cumby, R. E. and Obstfeld, M. (1981) A Note on Exchange - rate Expectations and Nominal Interest Differentials: A Test of the Fisher Hypothesis, Journal of Finance, 36(3), 697 - 703.

164. Cumby, R. E., Huizinga, J. and Obstfeld, M. (1983) Two - step Two - stage Least Squares Estimation in Models with Rational Expectations, Journal of Econometrics, 21, 333 - 355.

165. Curcio, R., Goodhart, C., Guillaume, D. and Payne, R. (1997) Do Technical Trading Rules Generate Profits. Conclusions from the Intra - day Foreign Exchange Market, International Journal of Finance and Economics, 2(4), 267 - 280.

166. Cuthbertson, K. (1991a) Modelling the Demand for Money in C. J. Green and D. T. Llewellyn (eds.), Surveys in Monetary Economics, Volume I, Monetary Theory and Policy, Blackwell, Oxford.

167. Cuthbertson, K. (199 lb) The Encompassing Implications of Feedforward Versus Feedback Mechanisms: A Reply to Hendry, Oxford Economic Papers, 43(2), 344 - 350.

168. Cuthbertson, K. (1996) The Expectations Hypothesis of the Term Structure: The UK Interbank Market, Economic Journal, 106(436), 578 - 592.

169. Cuthbertson, K. (1997) Microfoundations and the Demand for Money, Economic Journal, 107(443), 1186 - 1201.

170. Cuthbertson, K. and Bredin, D. (2000) The Expectations Hypothesis of the Term Structure: The Case of Ireland, Economic and Social Review, 31(3), 267 - 281.

171. Cuthbertson, K. and Bredin, D. (2001) Risk Premia and Long Rates in Ireland, Journal of Forecasting, 20. 91 - 403.

172. Cuthbertson, K. and Galindo, L. (1999) The Demand for Money in Mexico, Manchester School, 67(2), 154 - 166.

173. Cuthbertson, K. and Gasparro, D. (1993) The Determinants of Manufacturing Inventories in the UK, Economic Journal, 103(421), 1479.

174. Cuthbertson, K. and Hyde, S. (2002) Excess Volatility and Efficiency in French and German Stock Markets, Economic Modelling, 19(3), 3.

175. Cuthbertson, K. and Hyde, S. (2004) Resurrecting the C - CAPM: Empirical Evidence from France and Germany, Working Paper, School of Accounting and Finance, University of Manchester.

176. Cuthbertson, K. and Nitzsche, D. (2001a) Investments: Spot and Derivatives Markets, John Wiley, Chichester. Cuthbertson, K. and Nitzsche, D. (2001b) Financial Engineering: Derivatives and Risk Management, John Wiley, Chichester.

177. Cuthbertson, K. and Nitzsche, D. (2003) Long Rates, Risk Premia and the Over - reaction Hypothesis, Economic Modelling, 20(2), 417 - 435.

178. Cuthbertson, K., Hayes, S. and Nitzsche, D. (1996) The Behaviour of Certificate of Deposit Rates in the UK, Oxford Economic Papers, 48, 397 - 414.

179. Cuthbertson, K., Hayes, S. and Nitzsche, D. (1997) The Behaviour of UK Stock Prices and Returns: Is the Market Efficient? Economic Journal, 107(443), 986 - 1008.

180. Cuthbertson, K., Hayes, S. and Nitzsche, D. (1999) Market Segmentation and Stock Price Behaviour, Oxford Bulletin of Economics and Statistics, 61(2), 217 - 236.

181. Cuthbertson, K., Hayes, S. and Nitzsche, D. (2000) Are German Money Market Rates Well Behaved?

Journal of Economics, Dynamics and Control, 34, 347 - 380.

182. Cuthbertson, K. , Nitzsche, D. and O'Sullivan, N. (2004) UK Mutual Fund Performance: Skill or Luck?, Cass Business School Working Paper, mimeo.

183. Dacco, R. and Satchell, S. E. (1999) Why Do Regime Switching Models Forecast So Badly? Journal of Forecasting, 18, 1 - 16.

184. Dai, Q. and Singleton, K. (2003) Term Structure Dynamics in Theory and Reality, Review of Financial Studies, 16(3), 631 - 678.

185. Daniel, K. , Hirshleifer, D. and Subrahmanyam, A. (1998) Investor Psychology and Security Market Under and Overreactions, Journal of Finance, 53(6), 1839 - 1885.

186. Danielsson, J. and deVries, C. G. (1997) Extreme Returns, Tail Estimation and Value at Risk, Journal of Empirical Finance, 4, 241 - 257.

187. Day, T. E. and Lewis, C. M. (1992) Stock Market Volatility and the Information Content of Stock Index Options, Journal of Econometrics, 52, 267 - 287.

188. Deaton, A. and Paxson, C. (1994) Intertemporal Choice and Inequality, Journal of Political Economy, 102 (3), 437 - 467.

189. DeBondt, W. F. M. and Thaler, R. H. (1985) Does the Stock Market Overreact? Journal of Finance, 40 (3), 793 - 805.

190. DeBondt, W. F. M. and Thaler, R. H. (1987) Further Evidence on Investors Overreaction and Stock Market Seasonally, Journal of Finance, 42, 557 - 581.

191. DeBondt, W. F. M. and Thaler, R. H. (1989) Anomalies: A Mean - reverting Walk Down Wall Street, Journal of Economic Perspectives, 3(1), 189 - 202.

192. DeGrauwe, P. , Dewachter, H. and Embrechts, M. (1993) Exchange Rate Theory: Chaotic Models of Foreign Exchange Markets, Blackwell, Oxford.

193. deGregorio, J. , Giovanni, F. and Wolf, H. C. , Gordon, R. J. and deMenil, G. (1994) International Evidence on Tradables and Nontradables Inflation, European Economic Review, 38(6), 1225 - 1256.

194. DelGuercio, D. and Tkac, P. A. (2000) The Determinants of Flow of Funds of Managed Portfolios: Mutual Funds versus Pension Funds, Working paper, Lundquist College of Business, University of Oregon.

195. DeLong, B. J. , Shleifer, A. , Summers, L. H. and Waldmann, R. J. (1990a) Positive Feedback Investment Strategies and Destabilizing Rational Speculation, Journal of Finance, 45(2), 379 - 395.

196. DeLong, J. B. , Shleifer, A. , Summers, L. H. and Waldmann, R. J. (1990b) Noise Trader Risk in Financial Markets, Journal of Political Economy, 98(4), 703 - 738.

197. Dhrymes, P. T. , Friend, I. and Gultekin, N. B. (1984) A Critical Re - examination of the Empirical Evidence on the APT, Journal of Finance, 39(2), 323 - 346.

198. Diba, B. T. and Grossman, H. L. (1988) Explosive Rational Bubbles in Stock Prices? American Economic Review, 78(3), 520 - 530.

199. Dickey, D. A. and Fuller, W. A. (1979) Distribution of the Estimators for Autoregressive Time Series with a Unit Root, Journal of the American Statistical Association, 74(366), 427 - 431.

200. Diebold, F. X. (1988) Serial Correlation and the combination of forecasts, Journal of Business and Economic Statistics, 6, 105 - 111.

201. Diether, K. , Malloy, C. and Scherbina, A. (2002) Stock Prices and Differences of Opinion: Empirical Evidence that Stock Prices Reflect Optimism, Journal of Finance, 57, 2113 - 2141.

202. Dimson, E. , Marsh, P. and Staunton, M. (2002) Triumph of the Optimists: 101 Years of Global Investment Returns, Princeton University Press, Princeton, N. J.

203. Dooley, M. P. and Shafer, J. R. (1983) Analysis of Short - run Exchange Rate Behavior: March 1973 to November 1981, in T. Bigman and T. Taya (eds.), Exchange Rate and Trade Instability: Causes, Consequences

and Remedies, Harper Business, pp. 43 - 69.

204. Dornbusch, R. (1976) Expectations and Exchange Rate Dynamics, Journal of Political Economy, 84(6), 1161 - 1176.

205. Dornbusch, R. and Fischer, S. (1980) Exchange Rates and the Current Account, American Economic Review, 70(5), 960 - 971.

206. Dowd, K. (2000) Adjusting for Risk: An Improved Sharpe Ratio, International Review of Economics and Finance, 9, 209 - 220.

207. Driffell, J. and Sola, M. (1994) Testing the Term Structure of Interest - rates Using Stationary Vector Autoregression with Regime Switching, Journal of Economic Dynamics and Control, 18(3/4), 601 - 628.

208. Driffell, J. and Sola, M. (1998) Intrinsic Bubbles and Regime Switching, Journal of Monetary Economics, 42(2), 357 - 374.

209. Droms, W. and Walker, D. (1994) Investment Performance of International Mutual Funds, Journal of Financial Research, 17(1), 1 - 11.

210. Dumas, B. (1992) Dynamic Equilibrium and the Rural Exchange Rate in a Separated World, Review of Financial Studies, 5(2), 153 - 180.

211. Edelen, R. M. (1999) Investor Flows and the Assessed Performance of Open - end Mutual Funds, Journal of Financial Economics, 53, 439 - 466.

212. Elliott, G. and Ito, T. (1999) Heterogeneous Expectations and Tests of Efficiency in the Yen/Dollar Forward Exchange Rate Market, Journal of Monetary Economics, 43(2), 435 - 456.

213. Ellsberg, D. (1961) The Crude Analysis of Strategic Choice, American Economic Review, 51(2), 472 - 478.

214. Elton, E., Gruber, M. and Blake, C. (1996) The Persistence of Risk Adjusted Mutual Fund Performance, Journal of Business, 69(2), 133 - 157.

215. Elton, E., Gruber, M., Das, S. and Hlavka, M. (1993) Efficiency with Costly Information: A Reinterpretation of Evidence from Managed Portfolios, Review of Financial Studies, 6, 1 - 21.

216. Engle, C. (1993) Real Exchange Rates and Relative Prices? An Empirical Investigation, Journal of Monetary Economics, 32(1), 35 - 50.

217. Engle, C. (1996) The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence, Journal of Empirical Finance, 3(2), 123 - 192.

218. Engle, C. and Hamilton, J. D. (1990) Long Swings in the Dollar: Are They in the Data and Do Market Know it, American Economic Review, 80(1), 689 - 713.

219. Engle, C. and Rogers, J. H. (1996) How Wide is the Border, American Economic Review, 86, 1112 - 1125.

220. Engle, R. F. and Bollerslev, T. (1986) Modelling the Persistence of Conditional Variances, Econometric Review, 5(1), 1 - 50.

221. Engle, R. F. and Granger, C. W. J. (1987) Co - integration and Error Correction: Representation, Estimation, and Testing, Econometrica, 55(2), 251 - 276.

222. Engle, R. F. and Kroner, K. K. (1995) Multivariate Simultaneous Generalized ARCH, Econometric Theory, 11(1), 122 - 150.

223. Engle, R. F., Ito, T. and Lin, W. L. (1990) Meteor Showers or Heat Waves? Heteroscedastic Intra - daily Volatility in the Foreign Exchange Market, Econometrica, 58, 525 - 542.

224. Engle, R. F., Lilien, D. M. and Robins, R. P. (1987) Estimating Time Varying Risk Premia in the Term Structure: The ARCH - M Model, Econometrica, 55(2), 391 - 407.

225. Engsted, T. (1996) The Predictive Power of the Money Market Term Structure, International Journal of forecasting, 12, 289 - 295.

226. Engsted, T. (1998) Evaluating the Consumption - capital Asset Pricing Model Using Hansen - Jagannathan Bounds: Evidence from the UK, International Journal of Finance and Economics, 3, 291 - 302.

227. Engsted, T. and Tanggaard, C. (1995) The Predictive Power of Yield Spreads for Future Interest Rates: Evidence from the Danish Term Structure, Scandinavian Journal of Economics, 97(1), 145 - 159.

228. Engsted, T. C. (2002) Measures of Fit for Rational Expectations Models, Journal of Economic Surveys, 16 (3), 301 - 356.

229. Epstein, L. G. and Zin, S. E. (1989) Substitution, Risk Aversion and the Temporal Behavior of Consumption Growth and Asset Returns: A Theoretical Framework, Econometrica, 57(4), 937 - 969.

230. Epstein, L. G. and Zin, S. E. (1991) Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns - An Empirical Analysis, Journal of Political Economy, 99(2), 263 - 286.

231. Eun, C. S. and Resnick, B. G. (1997) International Equity Investment with Selective Hedging Strategies, Journal of International Financial Markets, Institutions and Money, 7(1), 21 - 42.

232. Evans, G. W. (1991) Pitfalls in Testing for Explosive Bubbles in Asset Prices, American Economic Review, 81(4), 922 - 930.

233. Evans, G. W. and Honkapohja, S. (2001) Learning and Expectations Macroeconomics, Princeton University Press, Princeton, N. J.

234. Evans, M. D. (1998) Dividend Variability and Stock Market Swings, Review of Economic Studies, 65, 711 - 740.

235. Evans, M. D. and Lyons, R. K. (2002) Order Flow and Exchange Rate Dynamics, Journal of Political Economy, 110(1), 170 - 180.

236. Fama, E. F. (1984) Forward and Spot Exchange Rates, Journal of Monetary Economics, 14, 319 - 338.

237. Fama, E. F. and French, K. R. (1988b) Dividend Yields and Expected Stock Returns, Journal of Financial Economics, 22, 3 - 25.

238. Fama, E. F. and French, K. R. (1993) Common Risk Factors in the Returns on Stocks and Bonds, Journal of Financial Economics, 33, 3 - 56.

239. Fama, E. F. and French, K. R. (1998) Value Versus Growth: The International Evidence, Journal of Finance, 51, 55 - 84.

240. Fama, E. F. and French, K. R. (1988a) Permanent and Temporary Components of Stock Prices, Journal of Political Economy, 96, 246 - 273.

241. Fama, E. F. and French, K. R. (1995) Size and Book - to - market Factors in Earnings and Returns, Journal of Finance, 50(1), 131 - 155.

242. Fama, E. F. and French, K. R. (1996) Multifactor Explanations of Asset Pricing Anomalies, Journal of Finance, 47, 426 - 465.

243. Fama, E. F. and MacBeth, J. (1973) Risk, Return and Equilibrium: Empirical Tests, Journal of Political Economy, 81(3), 607 - 636.

244. Fama, E. F. and MacBeth, J. D. (1974) Tests of the Multiperiod Two - parameter Model, Journal of Financial Economics, 1(1), 43 - 66.

245. Fama, E. F. and French, K. R. (2002) The Equity Premium, Journal of Finance, 57(2), 637 - 659.

246. Ferson, W. and Schadt, R. (1996) Measuring Fund Strategy and Performance in Changing Economic Conditions, Journal of Finance, 51(2), 425 - 462.

247. Fisher, M. and Gilles, C. (1998) Around and Around: The Expectations Hypothesis, Journal of Finance, 53(1), 365 - 386.

248. Flavin, M. A. (1983) Excess Volatility in the Financial Markets: A Reassessment of the Empirical Evidence, Journal of Political Economy, 91(6), 929 - 956.

249. Fletcher, J. (1995) An Examination of the Selectivity and Market Timing Performance of UK Unit

Trusts, Journal of Business Finance and Accounting, 22(1), 143 - 156.

250. Fletcher, J. (1997) An Examination of UK Unit Trust Performance Within the Arbitrage Pricing Framework, Review of Quantitative Finance and Accounting, 8, 91 - 107.

251. Fletcher, J. (1999) The Evaluation of the Performance of UK American Unit Trusts, International Review of Economics and Finance, 8(4), 455 - 466.

252. Flood, R. P. and Garber, P. M. (1980) Market Fundamentals Versus Price Level Bubbles: The First Tests, Journal of Political Economy, 88(4), 745 - 770.

253. Flood, R. P. and Garber, P. M. (1994) Speculative Bubbles, Speculative Attacks and Policy Switching, MIT Press, Cambridge, Mass.

254. Flood, R. P. and Rose, A. K. (1995) Fixing Exchange Rates: A Virtual Quest for Fundamentals, Journal of Monetary Economics, 36(1), 3 - 38.

255. Flood, R. P. and Rose, A. K. (1996) Fixes of the Forward Discount Puzzle, Review of Economic Studies, 40(1 - 2), 209 - 224.

256. Flood, R. P., Garber, P. M. and Scott, L. O. (1984) Multi - country Tests for Price - level Bubbles, Journal of Economic Dynamics and Control, 84(3), 329 - 340.

257. Flood, R. P., Hodrick, R. J. and Kaplan, P. (1994) An Evaluation of Recent Evidence on Stock Price Bubbles, in R. P. Flood and P. M. Garber (eds.), Speculative Bubbles, Speculative Attacks and Policy Switching, MIT Press, Cambridge, Mass., pp. 105 - 133.

258. Fortin, R. and Michelson, S. E. (1995) Are Load Mutual Funds Worth the Price, Journal of Investing, 4(3), 89 - 94.

259. Fortune, P. (1991) Stock Market Efficiency: An Autopsy? New England Economic Review, Federal Reserve Bank of Boston, March/April, 17 - 40.

260. Fox, C. and Tversky, A. (1995) Ambiguity Aversion and Comparative Ignorance, Quarterly Journal of Economics, 110(3), 585 - 604.

261. Frankel, J. A. (1979) On the Mark: A Theory of Floating Exchange Rates Based on Real Interest Differentials, American Economic Review, 69(4), 610 - 622.

262. Frankel, J. A. (1982a) A Test of Perfect Substitutability in the Foreign Exchange Market, Southern Economic Review, 49(2), 406 - 416.

263. Frankel, J. A. (1982b) In Search of the Exchange Risk Premium: A Six Currency Test Assuming Mean - variance Optimization, Journal of International Money and Finance, 1, 255 - 274.

264. Frankel, J. A. and Froot, K. A. (1986) The Dollar as an Irrational Speculative Bubble: The Tale of Fundamentalists and Chartists, Marcus Wallenberg Papers on International Finance, 1, 27 - 55.

265. Frankel, J. A. and Froot, K. A. (1987) Using Survey Data to Test Standard Propositions Regarding Exchange Rate Expectations, American Economic Review, 77(1), 133 - 153.

266. Frankel, J. A. and Froot, K. A. (1988) Chartist, Fundamentalists and the Demand for Dollars, Greek Economic Review, 10, 49 - 102.

267. Frankel, J. A. and Rose, A. K. (1995) Empirical Research on Nominal Exchange Rates, in G. Grossman and K. Rogoff (eds.), Handbook of International Economics, III, North Holland, Amsterdam, pp. 1689 - 1729.

268. Frankel, J. A. and Rose, A. K. (1996) A Panel Project on Purchasing Power Parity: Mean Reversion within and between Countries, Journal of International Economics, 40, 209 - 224.

269. Franses, P. H. and vanDijk, D. (1996) Forecasting Stock Market Volatility Using non - linear GARCH Models, Journal of Forecasting, 15, 229 - 235.

270. French, K. R. and Roll, R. (1986) Stock Return Variances: The Arrival of Information and the Reaction of Traders, Journal of Financial Economics, 17, 5 - 26.

271. French, K. R., Schwert, G. W. and Stamburgh, R. F. (1987) Expected Stock Returns and Volatility, Jour-

nal of Financial Economics, 19, 3 – 29.

272. Frenkel, J. A., Gylfason, T. and Helliwell, J. F. (1980) A Synthesis of Monetary and Keynesian Approaches to Short – run Balance – of – payments Theory, Economic Journal, 90, 582 – 592.

273. Frino, A. and Gallagher, D. (2001) Tracking S&P500 Index Funds, Journal of Portfolio Management, 28, 44 – 55.

274. Froot, K. and Rogoff, K. (1991) The EMS, the EMU and the Transition to a Common Currency, NBER Macroeconomics Annual, 6, 269 – 317.

275. Froot, K. A. and Dabora, E. (1999) How are Stock Prices Affected by the Location of Trade? Journal of Financial Economics, 53, 189 – 216.

276. Froot, K. A. and Frankel, J. A. (1989) Forward Discount Bias: Is it an Exchange Risk Premium? Quarterly Journal of Economics, 104(1), 139 – 161.

277. Froot, K. A. and Obstfeld, M. (1991) Intrinsic Bubbles: The Case of Stock Prices, American Economic Review, 81(5), 1189 – 1214.

278. Frost, P. and Savarino, J. (1988) For Better Performance Constrain Portfolio Weights, Journal of Portfolio Management, 14, 29 – 34.

279. Fung, W. and Hsieh, D. A. (1997) Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds, Review of Financial Studies, 10, 275 – 302.

280. Gali, J. (1994) Keeping up with the Joneses: Consumption Externalities, Portfolio Choice, and Asset Pricing, Journal of Money, Credit and Banking, 26(1), 1 – 8.

281. Gallagher, L. A. and Taylor, M. P. (2001) Risky Arbitrage, Limits of Arbitrage and Non – linear Adjustment in the Dividend Price Ratio, Economic Enquiry, 39(4), 524 – 537.

282. Gallant, A. R., Hansen, L. P. and Tauchen, G. E. (1990) Using Conditional Moments of Asset Payoffs to Infer the Volatility of Intertemporal Marginal Rates of Substitution, Journal of Econometrics, 45, 141 – 180.

283. Garber, P. M. (1990) Famous First Bubbles, Journal of Economic Perspectives, 4(2), 35 – 54.

284. Gebhardt, W. R., Hvidkjaer, S. and Swaminathan, B. (2001) Stock and Bond Market Interaction: Does Momentum Spillover?, Unpublished Working Paper, Cornel University.

285. Gemmill, G. and Thomas, D. C. (2002) Noise Trading, Costly Arbitrage and Asset Prices: Evidence from Closed End Funds, Journal of Finance, LVII(6), 2571 – 2594.

286. Genesove, D. and Mayer, C. (2001) Loss Aversion and Seller Behavior: Evidence from the Housing Market, Quarterly Journal of Economics, 116, 1233 – 1260.

287. Geotzmann, W., Ingersoll Jr, J. and Ivokvich, Z. (2000) Monthly Measurement of Daily Timers, Journal of Financial and Quantitative Analysis', 35, 257 – 290.

288. Gerlach, S. (2003) Interpreting the Term Structure of Interbank Rates in Hong Kong, Pacific – Basin Finance Journal, 11, 593 – 609.

289. Gertner, R. (1993) Game Shows and Economic Behaviour – Risk Taking on "Card Sharks", Quarterly Journal of Economics, 108(2), 507.

290. Gilboa, I. and Schmeidler, D. (1989) Maximum Expected Utility with Non – unique Prior, Journal of Mathematical Economics, 18, 141 – 153.

291. Gilles, C. and LeRoy, S. F. (1991) Econometric Aspects of the Variance – bounds Tests: A Survey, Review of Financial Studies, 4(4), 753 – 791.

292. Gilles, C. and LeRoy, S. F. (1992) Bubbles and Changes, International Economic Review, 33(2), 323 – 339.

293. Giovanni, A. and Jorion, P. (1987) Interest Rates and Risk Premia in the Stock Market and in the Foreign Exchange Market, Journal of International Money and Finance, 6, 107 – 124.

294. Giovannini, A. and Weil, P. (1989) Risk Aversion and Intertemporal Substitution in the Capital Asset

Pricing Model, NBER Working Paper No. 2824.

295. Glosten, L. R. and Milgrom, P. R. (1985) Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders, Journal of Financial Economics, 14, 71 - 100.

296. Glosten, R. , Jagannathan, R. and Runkle, D. E. (1993) On the Relation Between Expected Value and the Volatility of the Normal Excess Return on Stocks, Journal of Finance, 48(5), 1779 - 1801.

297. Goetzmann, W. N. and Ibbotson, R. (1994) Do Winners Repeat, Journal of Portfolio Management, 20 (2), 9 - 18.

298. Goetzmann, W. N. and Jorion, P. (1999) Re - emerging Markets, Journal of Financial and Quantitative Analysis, 34(1), 1 - 32.

299. Goldberg, P. K. and Knetter, M. M. (1997) Goods Prices and Exchange Rates: What Have We Learned? Journal of Economic Literature, XXXV(3), 1243 - 1272.

300. Gong, F. and Remolona, E. M. (1997) Two Factors Along the Yield Curve, Manchester School, 75, 1 - 31.

301. Goodhart, C. and Curcio, R. (1992). When Support/Resistance Levels are Broken, Can Profits be Made? Evidence From the Foreign Exchange Market, Discussion Paper 142, London School of Economics, Financial Markets Group.

302. Goodhart, C. E. and Taylor, M. P. (1992) Why Don't Individuals Speculate in Forward Foreign Exchange? Scottish Journal of Political Economy, 39, 1 - 13.

303. Goodman, S. H. (1979) Foreign Exchange Rate Forecasting Techniques: Implications for Business and Policy, Journal of Finance, 34(2), 415 - 427.

304. Goodman, S. H. (1980) Who's Better than the Toss of a Coin? Euromoney, 80 - 84.

305. Gordon, M. J. (1962) The Investment, Financing and Valuation of the Corporation. , Irwin, Homewood, Ill.

306. Goyal, A. and Welch, I. (1999) The Myth of Predictability: Does the Dividend Yield Forecast the Equity Premium? Working Paper, UCLA, Calif.

307. Granger, C. W. J. and Terasvirta, T. (1993) Modelling Non - linear Economic Relationships, Oxford University Press, Oxford.

308. Gregory, A. W. and Veall, M. R. (1985) Formulating Wald Tests of Nonlinear Restrictions, Econometrica, 53(6), 1465 - 1468.

309. Grilli, V. and Kaminsky, G. (1991) Nominal Exchange Rate Regimes and the Real Exchange Rate: Evidence from the United States and Great Britain, 1885 - 1986, Journal of Monetary Economics, 27(2), 191 - 212.

310. Grinblatt, M. and Titman, S. (1992) The Persistence of Mutual Fund Performance, Journal of Finance, 47 (5), 1977 - 1984.

311. Grinblatt, M. and Titman, S. (1994) A Study of Monthly Mutual Fund Returns and Performance Evaluation Techniques, Journal of Financial and Quantitative Analysis, 29(3), 419 - 444.

312. Groen, J. J. (2000) Long Horizon Predictability of Exchange Rates: Is It for Real? Empirical Economics, 24(3), 451 - 469.

313. Grinblatt, M. and Keloharju, M. (2001) How Distance, Language, and Culture Influence Stockholdings and Trades, Journal of Finance, 56, 1053 - 1073.

314. Grossman, S. J. and Shiller, R. J. (1981) The Determinants of the Variability of Stock Market Prices, American Economic Review, 71, 222 - 227.

315. Grossman, S. J. and Stiglitz, J. E. (1980) The Impossibility of Informationally Efficient Markets, American Economic Review, 66 246 - 253.

316. Gruber, M. (1996) Another Puzzle: The Growth in Actively Managed Mutual Funds, Journal of Finance, 51(3), 783 - 810.

317. Gul, F. (1991) A Theory of Disappointment Aversion, Econometrica, 59(3), 667 - 686.

318. Gylfason, T. and Helliwell, J. F. (1983) A Synthesis of Keynesian, Monetary and Portfolio Approaches to Flexible Exchange Rates, Economic Journal, 93(372), 820 - 831.

319. Haache, G. and Townend, J. (1981) Exchange Rates and Monetary Policy: Modelling Sterling's Effective Exchange Rate, 1972 - 80, in W. A. Eltis and P. J. N. Sinclair (eds.), The Money Supply and the Exchange Rate, Oxford University Press, Oxford, pp. 201 - 247.

320. Hakkio, C. S. (1981) Expectations and the Forward Exchange Rate, International Economic Review, 22, 383 - 417.

321. Hall, A. D., Anderson, H. M. and Granger, C. W. J. (1992) A Cointegration Analysis of Treasury Bill Yields, Review of Economic and Statistics, 74, 116 - 126.

322. Hall, S. G., Miles, D. K. and Taylor, M. P. (1989) Modelling Asset Prices with Time Varying Betas: Some Evidence from the London Stock Exchange, Manchester School, L ~ II(4), 340 - 356.

323. Hamilton, J. D. (1988) Rational - expectations Econometric Analysis of Changes in Regime, Journal of Economic Dynamics and Control, 12(3), 385 - 423.

324. Hamilton, J. D. (1990) Analysis of Time Series Subject to Changes in Regimes, Journal of Econometrics, 45(1/2), 39 - 70.

325. Hamilton, J. D. (1994) Time Series Analysis, Princeton University Press, Princeton, N. J.

326. Hannan, E. J. (1970) Multiple Time Series, John Wiley, New York.

327. Hansen, L. P. (1982) Large Sample Properties of Generalized Method of Moments Estimators, Econometrica, 50(4), 1029 - 1054.

328. Hansen, L. P. and Hodrick, R. J. (1980) Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis, Journal of Political Economy, 88(5), 829 - 853.

329. Hansen, L. P. and Hodrick, R. J. (1983) Risk Averse Speculation in the Forward Foreign Exchange Market: An Econometric Analysis of Linear Models, in J. A. Frenkel (ed.), Exchange Rates and International Macroeconomics, University of Chicago Press, pp. 113 - 142.

330. Hansen, L. P. and Jagannathan, R. (1991) Restrictions on Intertemporal Marginal Rates of Substitution Implied by Asset Returns, Journal of Political Economy, 99, 225 - 262.

331. Hansen, L. P. and Richard, S. F. (1987) The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models, Econometrica, 55(3), 687 - 613.

332. Hansen, L. P., Heaton, J. and Luttmer, E. (1995) Econometric Evaluation of Asset Pricing Models, The Review of Financial Studies, 8, 237 - 274.

333. Hansen, P. R. (2002) Structural Changes in the Cointegrated Vector Autoregressive Model, Journal of Econometrics, 114, 195 - 295.

334. Harris, R. D. (2001) The Expectations Hypothesis of the Term Structure and Time - varying Risk Premia: A Panel Data Approach, Oxford Bulletin of Economics and Statistics, 63(2), 233 - 246.

335. Harris, R. D. and Sanchez - Valle, R. (2000a) The Information Content of Lagged Equity and Bond

336. Yields', Economic Letters, 68, 179 - 184.

337. Harris, R. D. and Sanchez - Valle, R. (2000b) The Gilt - equity Yield Rates and the Predictability of UK and US Equity Returns, Journal of Business Finance and Accounting, 27(3), 333 - 357.

338. Hau, H., Killeen, W. and Moore, M. (2002) The Euro as an International Currency: Explaining Puzzling First Evidence from the Foreign Exchange Markets, Journal of International Money and Finance, 21, 351 - 383.

339. Hausman, J. A. (1978) Specification Tests in Econometrics, Econometrica, 46, 1251 - 1272.

340. He, H. and Modest, D. (1995) Market Frictions and Consumption Based Asset Pricing, Journal of Political Economy, 103, 94 - 117.

341. Heaton, J. (1995) An Empirical Investigation of Asset Pricing with Temporally Dependent Preference

Specifications, Econometrica, 63(3), 681 - 717.

342. Heaton, J. and Lucas, D. (1996) Evaluating the Effects of Incomplete Markets on Risk - sharing and asset Pricing, Journal of Political Economy, 103, 94 - 117.

343. Hendricks, D., Patel, J. and Zeckhauser, R. (1993) Hot Hands in Mutual Funds: Short Run Persistence of Performance, Journal of Finance, 48, 93 - 130.

344. Hendry, D. F. (1995) Dynamic Econometrics, Oxford University Press, Oxford.

345. Hensen, L. P. and Richard, S. F. (1987) The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models, Econometrica, 55 587 - 613.

346. Hoare Govett (1991) UK Market Prospects for the Year Ahead, Equity Market Strategy.

347. Hodges, S. (1998) A Generalization of the Sharpe Ratio and its Application to Valuation Bounds and Risk Measures, FORC Preprint 98/88, April, University of Warwick.

348. Hodrick, R. J. (1987) The Empirical Evidence of the Efficiency of Forward and Future Foreign Exchange Markets, Harwood, Chur.

349. Hodrick, R. J. (1989) Risk Uncertainty and Exchange Rates, Journal of Monetary Economics, 23, 433 - 459.

350. Hogg, R. V. and Tanis, E. A. (1993) Probability and Statistical Inference, Macmillan, London.

351. Hong, H. and Stein, J. C. (1999) A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets, Journal of Finance, 54(6), 2143 - 2184.

352. Hong, H., Lim, T. and Stein, J. C. (2000) Bad News Travels Slowly: Size, Analyst Coverage and the Profitability of Momentum Strategies, Journal of Finance, 55(1), 265 - 295.

353. Huang, R. D. (1989) An Analysis of Intertemporal Pricing for Forward Exchange Contracts, Journal of Finance, 44, 181 - 194.

354. Huberman, G. (2001) Familiarity Breeds Investments, Review of Financial Studies, 14, 659 - 680.

355. Ibbotson Associates (2001) Stocks, Bonds, Bills and Inflation Year book, Ibbotson Associates, Chicago.

356. Ikenberry, D., Lakonishok, J. and Vermaelen, T. (1995) Market Underreaction to Open Market Share Repurchases, Journal of Financial Economics, 48, 181 - 208.

357. Ippolito, R. A. (1989) Efficiency with Costly Information: A Study of Mutual Fund Performance 1965 - 1984, Quarterly Journal of Economics, 104(1), 1 - 23.

358. Isard, P. (1978) Exchange Rate Determination: A Survey of Popular Views and Recent Models, Princeton Studies in International Finance, No. 42.

359. Ito, T. (1990) Foreign Exchange Rate Expectations: Micro Survey Data, American Economic Review, 80 (3), 434 - 449.

360. Ito, T., Engle, R. F. and Lin, W. L. (1992) Where Does the Meteor Shower Come From? The Role of Stochastic Policy Coordination, Journal of International Economics, 32, 221 - 240.

361. Jackson, P., Maude, D. J. and Perraudin, W. (1997) Bank Capital and Value - at - risk, Journal of Derivatives, 4, 73 - 90.

362. Jagannathan, R., McGatten, E. R. and Scherbina, A. (2001) The Declining US Equity Premium, NBER Working Paper No. 8172.

363. Jegadeesh, N. and Pennacchi, G. (1996) The Behaviour of Interest Rates Implied by the Term Structure of Eurodollar Futures, Journal of Money, Credit and Banking, 28(3), 47 - 70.

364. Jegadeesh, N. and Titman, S. (1993) Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, Journal of Finance, 48(1), 56 - 91.

365. Jegadeesh, N. and Titman, S. (2001) Profitability of Momentum Strategies and Evaluation of Alternative Explanations, Journal of Finance, 56, 699 - 720.

366. Jensen, M. (1968) The Performance of Mutual Funds in the Period 1945 - 1964, Journal of Finance, 23,

389 - 416.

367. Jensen, M. C. (1978) Some Anomalous Evidence Regarding Market Efficiency, Journal of Financial Economics, 6(2), 95 - 101.

368. Jensen, M. C. (1986) Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers, American Economic Review, 76(2), 323 - 332.

369. Jiang, W. (2003) A Nonparametric Test of Market Timing, Journal of Empirical Finance, 10, 399 - 425.

370. Jobson, J. D. and Korkie, B. (1984) On the Jensen Measure and Marginal Improvements in Portfolio Performance: A Note, Journal of Finance, 39, 245 - 252.

371. Jobson, J. D. and Korkie, B. M. (1980) Estimation for Markowitz Efficient Portfolios, Journal of the American Statistical Association, 75, 544 - 554.

372. Jobson, J. D. and Korkie, B. M. (1981) Putting Markowitz Theory to Work, Journal of Portfolio Management, 7, 70 - 74.

373. Johansen, S. (1988) Statistical Analysis of Cointegration Vectors, Journal of Economic Dynamics and Control, 12, 231 - 254.

374. Jorion, P. (2003) The Long - term Risks of Global Stock Markets, Graduate School of Management, University of California at Irvine, mimeo.

375. Kahneman, D. and Tversky, A. (1979) Prospect Theory: An Analysis of Decision under Risk, Econometrica, 47, 263 - 291.

376. Kaminsky, G. (1993) Is there a Peso Problem? Evidence from the Dollar/Pound Exchange Rate, 1976 - 1987, American Economic Review, 83(3), 450 - 472.

377. Kaminsky, G. and Peruga, R. (1990) Can a Time Varying Risk Premium Explain Excess Returns in the Forward Market for Foreign Exchange? Journal of International Economics, 28, 47 - 70.

378. Kandel, S. and Stamburgh, R. F. (1996) On the Predictability of Stock Returns: An Asset Allocation Perspective, Journal of Finance, 51, 385 - 424.

379. Keane, S. M. (1983) Stock Market Efficiency: Theory, Evidence and Implications, Philip Allan, Oxford.

380. Keim, D. B. and Stamburgh, R. F. (1986) Predicting Returns in the Stock and Bond Markets, Journal of Financial Economics, 17, 357 - 390.

381. Keynes, J. M. (1936) The General Theory of Employment, Interest, and Money, Harcourt, Brace and World, New York.

382. Kilian, L. and Taylor, M. P. (2003) Why is it so Difficult to Beat the Random Walk Forecast of Exchange Rates? Journal of International Economics, 60(1), 85 - 107.

383. Kim, T. S. and Omberg, E. (1996) Dynamic Nonmyopic Portfolio Behavior, Review of Economic Studies, 9(1), 141 - 161.

384. Kirman, A. P. (1991) Epidemics of Opinion and Speculative Bubbles in Financial Markets, in M. Taylor (ed.), Money and Financial Markets, Macmillan, London.

385. Kirman, A. P. (1993a) Ants, Rationality, and Recruitment, Quarterly Journal of Economics, 108(1), 137 - 156.

386. Kirman, A. P. (1993b) Testing for Bubbles, European University Institute, Florence, Ala.

387. Kleidon, A. W. (1986) Variance Bounds Tests and Stock Price Valuation Models, Journal of Political Economy, 94, 953 - 1001.

388. Knetter, M. M. (1989) Price Discrimination by US and German Exporters, American Economic Review, 79, 198 - 210.

389. Knetter, M. M. (1993) International Comparisons of Price - to - market Behaviour, American Economic Review, 83, 374 - 486.

390. Kocherlakota, N. R. (1996) The Equity Premium: It's Still a Puzzle, Journal of Economic Literature,

XXXI ~ ,42 - 71.

391. Koedijk,K. G. and Ott,M. (1987) Risk Aversion,Efficient Markets and the Forward Exchange Rate, Review,Federal Reserve Bank of St. Louis,69(10),5 - 13.

392. Kosowski,R. ,Timmermann,A. ,Wermers,R. and White,H. (2003) Can Mutual Fund Stars Really Pick Stocks? New Evidence from a Boostrap Analysis,Discussion Paper,University of Southern California,San Diego, Calif.

393. Krugman,P. R. (1978) Pricing to Market When the Exchange Rate Changes, in S. W. Arndt and J. D. Richardson (eds.),Real - financial Linkages Among Open Economies,MIT Press,Cambridge,Mass,pp. 49 - 70.

394. Krugman,P. R. (1990),Equilibrium Exchange Rates in International Monetary Policy Co - ordination and Exchange Rate Fluctuations,in W. H. Branson,J. A. Frankel and M. Goldstein (eds.),University of Chicago Press,Chicago,pp. 159 - 187.

395. Kugler,P. (1988) An Empirical Note on the Term Structure and Interest Rate Stabilization Policies, Quarterly Journal of Economics,1113(4),789 - 792.

396. Kugler,P. (2002) The Term Premium,Time Varying Interest Rate Volatility and Central Bank Policy Reaction,Economics Letters,76,311 - 316.

397. Kyle,A. and Xiong,W. (2001) Contagion as a Wealth Effect,Journal of Finance,56,1401 - 1443.

398. La Porta,R. ,Lakonishok,J. ,Shleifer,A. and Vishny,R. (1997) Good News for Value Stocks: Further Evidence on Market Efficiency,Journal of Finance,52(2),2.

399. Lakonishok,J. ,Shleifer,A. and Vishny,R. W. (1994) Contrarian Investment,Extrapolation and Risk, Journal of Finance,49(5),5.

400. Lamoureux,C. G. and Lastrapes,W. D. (1990a) Heteroscedasticity in Stock Return Data: Volume Versus GARCH Effects,Journal of Finance,45(1),221 - 229.

401. Lamoureux,C. G. and Lastrapes,W. D. (1990b) Persistence in Variance,Structural Change and the GARCH Model,Journal of Business and Economic Statistics,8,225 - 234.

402. Lane,P. R. (2001) The New Open Economy Macroeconomics: A Survey,Journal of International Economics,54(2),235 - 266.

403. Lasfer,M. A. ,Melnik,A. and Thomas,D. C. (2003) Short - term Reaction of Stock Markets in Stressful Circumstances,Journal of Banking and Finance,27,1959 - 1977.

404. Lee,C. M. C. ,Shleifer,A. and Thaler,R. (1991) Investor Sentiment and the Closed - end Fund Puzzle, Journal of Finance,46(1),75 - 110.

405. Lee,C. M. C. ,Shleifer,A. and Thaler,R. H. (1990) Closed - end Mutual Funds,Journal of Economic Perspectives,4(4),153 - 164.

406. Lee,T. H. ,White,H. and Granger,C. W. J. (1993) Testing for Neglected Nonlinearity in Time Series Models - A comparison of Neutral Network Methods and Alternative Tests,Journal of Econometrics,56(3),269 - 290.

407. Lee,W. Y. ,Jiang,C. X. and Indro,D. C. (2002) Stock Market Volatility,Excess Returns and the Role of Investor Sentiment,Journal of Banking and Finance,26,2277 - 2299.

408. Leger,L. (1997) UK Investment Trusts: Performance,Timing and Selectivity,Applied Economics Letters,4,207 - 210.

409. LeRoy,S. F. and Parke,W. R. (1992) Stock Market Volatility: Tests Based on the Geometric Random Walk,American Economic Review,82(4),981 - 992.

410. LeRoy,S. F. and Porter,R. D. (1981) The Present - value Relation: Tests Based on Implied Variance Bounds,Econometrica,49,555 - 574.

411. Lesmond,D. A. ,Schill,M. J. and Chunsheng,Z. (2004) The Illusory Nature of Momentum Profits,Jour-

nal of Financial Economics,71,349 - 380.

412. Lettau,M. and Ludvigson,S. (2001a) Consumption,Aggregate Wealth and Expected Stock Returns,Journal of Finance, $6,815 - 847.

413. Lettau,M. and Ludvigson,S. (2001b) Resurrecting the (C)CAPM: A Cross - sectional Test When Risk Premia are Time - varying,Journal of Political Economy,109(6),1238 - 1287.

414. Levich,R. M. (1980) Analysing the Accuracy of Foreign Exchange Advisory Services: Theory and Evidence,Chapter 5 in R. M. Levich and Wihlborg (eds.),Exchange Risk and Exposure,Lexington Books.

415. Levich,R. M. and Thomas,L. R. (1993) The Significance of Technical Trading Rule Profits in the

416. Foreign Exchange Market: A Bootstrap Approach,Journal of International Money and Finance,12,451 - 474.

417. Levy,E. and Nobay,R. (1986) The Speculative Efficiency Hypothesis: A Bivariate Analysis,Economic Journal,96,109 - 121,Conference Supplement.

418. Lewellen,J. W. (2002) Momentum and Autocorrelation in Stock Returns,Review of Financial Studies, 15,533 - 563.

419. Lewellen,J. and Shanken,J. (2002) Learning,Asset - pricing Tests and Market Efficiency,Journal of Finance,57(3),1113 - 1145.

420. Lewis, K. K. (1988) Testing the Portfolio Balance Model: A Multi - lateral Approach, Journal of International Economics,24,109 - 127.

421. Lewis,K. K. (1989a) Can Learning Affect Exchange Rate Behaviour? The Case of the Dollar in the Early 1980's,Journal of Monetary Economics,23(1),79 - 100.

422. Lewis,K. K. (1989b) Changing Beliefs and Systematic Rational Forecast Errors with Evidence from Foreign Exchange,American Economic Review,79(4),621 - 636.

423. Lewis,K. K. (1990) The Behaviour of Eurocurrency Returns Across Different Holding Periods and Regimes,Journal of Finance,45,1211 - 1236.

424. Lewis,K. K. (1991) Was There a "Peso Problem" in the U. S. Term Structure of Interest Rates - 1979 - 1983,International Economic Review,32(1),159 - 173.

425. Lewis,K. K. (1995) Are Foreign Exchange Rate Intervention and Monetary Policy Related,and Does it Really Matter? Journal of Business,68(2),185 - 214.

426. Li,H. and Maddala,G. S. (1997) Bootstrapping Cointegration Regressions,Journal of Econometrics,80, 297 - 318.

427. Lin,C. F. J. and Terasvirta,T. (1994) Testing the constancy of Regression Parameters Against Continuous Structural Change,Journal of Econometrics,62(2),211 - 228.

428. Lintner,J. (1971) The Aggregation of Investors' Diverse Judgements and Preferences in Purely Competitive Security Markets,Journal of Finance and Quantitative Analysis,4(4),347 - 450.

429. Liu,H. and Loewenstein,M. (2002) Optimal Portfolio Selection with Transaction Costs and Finite Horizons,Review of Financial Studies,15,805 - 835.

430. Lo,A. W. and MacKinlay,A. C. (1988) Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test,Review of Financial Studies,1,41 - 66.

431. Longstaff,F. A. (2000a) The Term Structure of Very Short - term Rates: New Evidence for the Expectations Hypothesis,Journal of Financial Economics,58,397 - 415.

432. Longstaff,F. A. (2000b) Arbitrage and the Expectations Hypothesis,Journal of Finance,55,989 - 994.

433. Lothian,J. R. and Taylor,M. P. (1996) Real Exchange Rate Behaviour: The Recent Float from the Perspective of the Past Two Centuries,Journal of Political Economy,104,488 - 510.

434. Lothian,J. R. and Taylor,M. P. (2000) Purchasing Power Parity over Two Centuries: Strengthening the Case for Real Exchange Rate Stability: A Reply to Coddington and Liang,Journal of International Money and

Finance, 19, 759 - 764.

435. Loughran, T. and Ritter, J. R. (1995) The New Issues Puzzle, Journal of Finance, 50, 23 - 51.

436. Loughran, T. and Ritter, J. R. (2000) Uniformly Least Powerful Tests of Market Efficiency, Journal of Financial Economics, 55(3), 361 - 389.

437. Lucas, R. E. (1978) Asset Prices in an Exchange Economy, Econometrica, 46, 1426 - 1446.

438. Lunde, A., Timmermann, A. and Blake, D. (1999) The Hazards of Mutual Fund Underperformance: A Cox Regression Analysis, Journal of Empirical Finance, 6, 121 - 152.

439. Luttmer, E. and Nishiotis, G. (1995) Joint Restrictions on Preferences with Non - traded Goods and Partially Integrated Financial Markets, Kellogg Graduate School of Management, Northwestern University, Evanston, Ill.

440. Lyons, R. K. (1995) Tests of Microstructural Hypotheses in the Foreign Exchange Market, Journal of Financial Economics, 39, 321 - 351.

441. Lyons, R. K. (1996) Foreign Exchange Volume: Sound and Fury Signifying Nothing?, in J. A. Frankel, G. Giampaolo, and G. Alberto, (eds.), The Microstructure of Foreign Exchange Markets, University of Chicago Press, Chicago, pp 183 - 205.

442. Lyons, R. K. (1998) Profits and Position Control: A Week of FX Dealing, Journal of International Money and Finance, 17, 97 - 115.

443. Lyons, R. K. (2001) The Microstructure Approach to Exchange Rates, MIT Press, Cambridge, Mass.

444. MacDonald, R. and Marsh, I. W. (1996) Foreign Exchange Forecasters are Heterogeneous: Confirmation and Consequences, Journal of International Money and Finance, 15, 665 - 685.

445. MacDonald, R. and Power, D. (1995) Stock Prices, Dividends and Retention: Long - Run Relationships and Short - Run Dynamics, Journal of Empirical Finance, 2, 135 - 151.

446. MacDonald, R. and Speight, A. E. H. (1991) The Term Structure of Interest Rates under Rational Expectations: Some International Evidence, Applied Financial Economics, 1, 211 - 221.

447. MacDonald, R. and Taylor, M. P. (1993) The Monetary Approach to the Exchange Rate: Rational Expectations, Long Run Equilibrium and Forecasting, IMF Staff Papers, 40(1), 89 - 107.

448. MacDonald, R. and Torrance, T. S. (1988) On Risk, Rationality and Excessive Speculation in the

449. Deutschemark United States Dollar Exchange Market: Some Evidence Using Survey Data, Oxford Bulletin of Economics and Statistics, 50(2), 107 - 123.

450. Maenhout, P. J. (2001) Robust Portfolio Rules, Hedging and Asset Pricing, Working Paper, Finance Department, INSEAD, France.

451. Mahoney, J. (1996) Forecasting Biases in Value - at - risk Estimations: Evidence from Foreign Exchange and Global Equity Portfolios, Working Paper, Federal Reserve Bank of New York, September.

452. Malkiel, B. G. (1977) The Valuation of Closed - end Investment - company Shares, Journal of Finance, 32(3), 847 - 859.

453. Malkiel, B. G. (1995) Returns from Investing in Equity Mutual Funds 1971 to 1991, Journal of Finance, 50(2), 549 - 572.

454. Mankiw, N. G. and Miron, J. A. (1986) The Changing Behavior of the Term Structure of Interest Rates, Quarterly Journal of Economics, 101(2), 211 - 228.

455. Mankiw, N. G. and Shapiro, M. D. (1986) Risk and Return: Consumption Beta Versus Market Beta, Review of Economics and Statistics, 68(3), 452 - 459.

456. Mankiw, N. G. and Summers, L. H. (1984) Do Long - term Interest Rates Overreact to Short Term Interest Rates, Brookings Papers on Economic Activity, 1, 223 - 242.

457. Mankiw, N. G. and Zeldes, S. P. (1991) The Consumption of Stockholders and Non - stockholders, Journal of Financial Economics, 29(1), 97 - 112.

458. Mankiw, N. G., Romer, D. and Shapiro, M. D. (1985) An Unbiased Reexamination of Stock Market Vola-

tility, Journal of Finance, 40, 677 – 687.

459. Mankiw, N. G., Romer, D. and Shapiro, M. D. (1991) Stock Market Forecastability and Volatility: A Statistical Appraisal, Review of Economic Studies, 58, 455 – 477.

460. Marcet, A. and Sargent, T. J. (1989) Convergence of Least Squares Learning Mechanisms in Self –

461. referential Linear Stochastic Models, Journal of Economic Theory, 48(2), 337 – 368.

462. Mark, N. C. (1985) On Time Varying Risk Premia in the Foreign Exchange Market: An Econometric Analysis, Journal of Monetary Economics, 16, 3 – 18.

463. Mark, N. C. (1988) Time Varying Betas and Risk Premia in the Pricing of Forward Foreign Exchange Contracts, Journal of Financial Economics, 22, 335 – 354.

464. Mark, N. C. (1995) Exchange Rates and Fundamentals: Evidence on Long Horizon Predictability, American Economic Review, 85(1), 201 – 218.

465. Mark, N. C. and Sul, D. (2001) Nominal Exchange Rates and Monetary Fundamentals: Evidence from a Small Post Bretton Woods Panel, Journal of International Economics, 53, 29 – 52.

466. Mark, N. C. and Wu, Y. (1998) Rethinking Deviations from Uncovered Interest Parity: The Role of Covariance Risk and Noise, Economic Journal, 108(451), 1686 – 1706.

467. Marsh, T. A. and Merton, R. C. (1986) Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices, American Economic Review, 711, 483 – 498.

468. McCallum, B. T. (1994) A Reconsideration of the Uncovered Interest Parity Relationship, Journal of Monetary Economics, 33, 105 – 132.

469. McCulloch, J. H. (1971) Measuring the Term Structure of Interest Rates, Journal of Business, 44, 19 – 31.

470. McCulloch, J. H. (1990) U. S. Government Term Structure Data, in B. Friedman and F. Hahn (eds.), The Handbook of Monetary Economics, North Holland, Amsterdam, The Netherlands.

471. McCulloch, J. H. (1993) A Reexamination of Traditional Hypotheses about the Term Structure – A Comment, Journal of Finance, 48(2), 779.

472. McElroy, M. B., Burmeister, E. and Wall, K. D. (1985) Two Estimators for the APT Model When Factors are Measured, Economic Letters, 19, 271 – 275.

473. McFarland, J. W., McMahon, P. C. and Ngama, Y. (1994) Forward Exchange Rates and Expectations during the 1920s: A Re – examination of the Evidence, Journal of International Money and Finance, 13, 627 – 636.

474. Meese, E. (1990) Currency Fluctuation in the Post – Bretton Woods Era, Journal of Economic Perspectives, 4(1), 117 – 134.

475. Meese, R. A. (1986) Testing for Bubbles in Exchange Markets: A Case of Sparkling Rates, Journal of Political Economy, 94(2), 345 – 373.

476. Meese, R. A. and Rogoff, K. (1983) Empirical Exchange Rate Models of the Seventies: Do they Fit Out of Sample? Journal of International Economics, 14, 3 – 24.

477. Mehra, R. and Prescott, E. C. (1985) The Equity Premium: A Puzzle, Journal of Monetary Economics, 15, 145 – 161.

478. Menkhoff, L. (1998) The Noise Trading Approach – Questionnaire Evidence from Foreign Exchange, Journal of International Money and Finance, 17, 547 – 564.

479. Merton, R. C. and Henriksson, R. (1981) On Market Timing and Investment Performance II: Statistical Procedures for Evaluating Forecasting Skills, Journal of Business, 54, 513 – 533.

480. Merton, R. C. (1969) Lifetime Portfolio Selection under Uncertainty. The Continuous – time Case, Review of Economics and Statistics, 51, 247 – 257.

481. Merton, R. C. (1971) Optimum Consumption and Portfolio Rules in a Continuous Time Model, Journal of Economic Theory, 3, 373 – 413.

482. Merton, R. C. (1973) An Intertemporal Capital Asset Pricing Model, Econometrica, 41, 867 – 887.

483. Merton, R. C. (1980) On Estimating the Expected Return on the Market, Journal of Financial Economics, 8, 323 – 361.

484. Merton, R. C. (1987) On the Current State of the Stock Market Rationality Hypothesis, in S. Fischer (ed.), Macroeconomics and Finance: Essays in Honor of Franco Modigliani, MIT Press, Cambridge, Mass.

485. Michael, P., Nobay, A. R. and Peel, D. A. (1997) Transactions Costs and Non – Linear Adjustment in Real Exchange Rates: An Empirical Investigation, Journal of Political Economy, 105, 862 – 879.

486. Michaely, R., Thaler, R. H. and Womack, K. L. (1995) Price Reactions to Dividend Initiations and Omissions: Overreaction or Draft?, Journal of Finance, 50, 573 – 608.

487. Miles, D. and Timmermann, A. (1996) Variation in Expected Stock Returns: Evidence on the Pricing of Equities from a Cross – section of UK Companies, Economica, 63(251), 369 – 382.

488. Modjtahedi, B. (1991) Multiple Maturities & Time Varying Risk Premia in Forward Exchange Markets: An Econometric Analysis, Journal of International Economics, 30, 69 – 86.

489. Moore, M. J. and Roche, M. J. (2002) Volatility and Persistent Real Exchange Rates without Sticky Prices, Quarterly Financial Economics.

490. Moore, M. J. and Roche, M. J. (2002) Less of a Puzzle: A New Look at the Forward Forex Market, Journal of International Economics, 58(2), 387 – 411.

491. Morey, M. R. (2003) Should You Carry the Load? A Comprehensive Analysis of Load and No – load Mutual Fund Out – of – sample Performance, Journal of Banking and Finance, 27, 1245 – 1271.

492. Moskowitz, T. and Grinblatt, M. (1999) Do Industries Explain Momentum? Journal of Finance, 54, 1249 – 1290.

493. Muth, J. F. (1961) Rational Expectations and the Theory of Price Movements, Econometrica, 29(3), 315 – 335.

494. Myers, S. and Majluf, N. (1984) Corporate Financing and Investment Decisions When Firms Have Information Investors Do Not Have, Journal of Financial Economics, 187 – 221.

495. Neely, C. J. and Weller, P. A. (2003) Intraday Technical Trading in the Foreign Exchange Market, Journal of International Money and Finance, 22, 223 – 237.

496. Neely, C. J., Weller, P. A. and Dittmar, R. (1997) Is Technical Analysis in the Foreign Exchange Market Profitable? A Genetic Programming Approach, Journal of Financial and Quantitative Analysis, 32, 405 – 426.

497. Nelson, D. B. (1991) Conditional Heteroscedasticity in Asset Returns, Econometrica, 59(2), 343 – 370.

498. Newey, W. K. and McFadden, D. L. (1994) Large Sample Estimation and Hypothesis Testing, in R. F. Engle and D. L. McFadden (eds.), Handbook of Econometrics: Volume 4, Elsevier Science, Amsterdam, The Netherlands, pp. 2111 – 2245.

499. Newey, W. K. and West, K. D. (1987) A Simple, Positive Semi – definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica, 55(3), 703 – 708.

500. Ng, D. T. (2004) The International CAPM When Expected Returns are Time Varying, Journal of International Money and Finance, 23, 189 – 230.

501. Obstfeld, M. and Rogoff, K. (2000) New Directions for Stochastic Open Economy Models, Journal of International Economics, 20(1), 1.

502. Obstfeld, M. and Taylor, A. M. (1997) Nonlinear Aspects of Goods – market Arbitrage and Adjustment: Heckscher's Commodity Points Revisited, Journal of the Japanese and International Economies, 11, 441 – 479.

503. Odean, T. (1998) Are Investors Reluctant to Realize their Losses? Journal of Finance, 53, 1775 – 1798.

504. Odean, T. (1999) Do Investors Trade Too Much? American Economic Review, 89, 1279 – 1298.

505. Pagan, A. R. and Ullah, A. (1988) The Econometric Analysis of Models with Risk Terms, Journal of Applied Econometrics, 3, 87 – 105.

506. Parsley, D. C. and Wei, S. J. (1996) Convergence to the Law of One Price without Trade Barriers or Currency Fluctuations, Quarterly Journal of Economics, 111, 1211 - 1136.

507. Pastor, L. (2000) Portfolio Selection and Asset Pricing Models, Journal of Finance, LV(1), 179 - 223.

508. Patterson, K. (2000) An Introduction to Applied Econometrics, Palgrave.

509. Peel, D. A. and Taylor, M. P. (2002) Covered Interest Arbitrage in the Inter - war Period and the Keynes - Einzig Conjecture, Journal of Money, Banking and Credit, 34(1), 51 - 75.

510. Pesando, J. E. (1983) On Expectations, Term Premiums and the Volatility of Long - term Interest Rates, Journal of Monetary Economics, 12(3), 467 - 474.

511. Pesaran, M. H. (1987) The Limits to Rational Expectations, Blackwell, Oxford.

512. Pesaran, M. H. and Timmermann, A. (1994) Forecasting Stock Returns: An Examination of Stock Market Trading in the Presence of Transaction Costs, Journal of Forecasting, 13(4), 335 - 367.

513. Pesaran, M. H. and Timmermann, A. (1995) Predictability of Stock Returns: Robustness and Economic Significance, Journal of Finance, 50(4), 1201 - 1228.

514. Pesaran, M. H. and Timmermann, A. (2000) A Recursive Modelling Approach to Predicting UK Stock Returns, Economic Journal, 110(460), 159 - 191.

515. Phillips, P. C. B. and Hansen, B. E. (1990) Statistical Inference in Instrumental Variables Regression with I(1) Processes, Review of Economic Studies, 57, 99 - 125.

516. Phillips, P. C. B. and Perron, P. (1988) Testing for a Unit Root in Time Series Regression, Biometrika, 75 (2), 335 - 346.

517. Piazzesi, M. (2002) Affine Term Structure Models, Handbook of Financial Econometrics, North Holland.

518. Politis, D. N. and Romano, J. P. (1994) The Stationary Bootstrap, Journal of the American Statistical Association, 89, 1303 - 1313.

519. Poterba, J. M. and Summers, L. H. (1986) The Persistence of Volatility and Stock Market Fluctuations, American Economic Review, 76(5), 1142 - 1151.

520. Poterba, J. M. and Summers, L. H. (1988) Mean Reversion in Stock Prices: Evidence and Implications, Journal of Financial Economics, 22, 26 - 59.

521. Prat, G. and Uctum, R. (1994) Formation des Anticipations de Change: les Enseignements des Enqueteeres du Consensus Forecasts, AFFI 16th Annual Meeting, Paris, December.

522. Psaradakis, Z. (2001) On Bootstrap Inference in Cointegrating Regressions, Economics Letters, 72, 1 - 10.

523. Quigley, G. and Sinquefield, R. (1999) The Performance of UK Equity Unit Trusts, Report by Dimensional Fund Advisors for Institute for Fiduciary Education.

524. Rabin, M. (1998) Psychology and Economics, Journal of Economic Literature, XXXVI(1), 11 - 46.

525. Rabin, M. (2000) Risk Aversion and Expected Utility Theory: A Calibration Theorem, Econometrica, 68 (5), 5.

526. Rabin, M. and Thaler, R. H. (2001) Anomalies: Risk Aversion, Journal of Economic Perspectives, 15 (1), 219 - 232.

527. Rapach, D. E. and Wohar, M. E. (2002) Testing the Monetary Model of Exchange Rate Determination: New Evidence from a Century of Data, Journal of International Economics, 58(1), 359 - 386.

528. Reinganum, M. R. (1982) A Direct Test of Roll's Conjecture on the Firm Size Effect, Journal of Finance, 37(1), 27 - 35.

529. Reinganum, M. R. (1983) The Anomalous Stock Market Behavior of Small Firms in January: Empirical Tests for Tax - loss Selling Effects, Journal of Financial Economics, 12(1), 89 - 104.

530. Remola, E. M., Kleiman, P. and Gruenstein, D. (1997) Market Returns and Mutual Fund Flows, Economic Policy Review, Federal Reserve Bank of New York, July.

531. Remolona, E. M. , Wickens, M. R. and Gong, F. F. (1998) What was the Market's View of UK Monetary Policy? Estimating Inflation Risk and Expected Inflation with Indexed Bonds, Federal Reserve Bank of New York Discussion Paper.

532. Richards, A. (1995) Comovements in National Stock Market Returns: Evidence of Predictability but Not Cointegration, Journal of Monetary Economics, 36, 631 - 654.

533. Richards, A. (1997) Winner - loser Reversals in National Stock Market Indices: Can They Be Explained? Journal of Finance, 51, 2129 - 2144.

534. Robinson, P. M. (1994) Efficient Tests of Nonstationary Hypothesis, Journal of the American Statistical Association, 89, 1420 - 1437.

535. Rogers, J. H. and Jenkins, M. (1995) Haircuts or Hysteresis? Sources of Movements in Real Exchange Rates, Journal of International Economics, 38, 339 - 360.

536. Rogoff, K. (1984) On the Effects of Sterilized Intervention: An Analysis of Weekly Data, Journal of Monetary Economics, 14, 133 - 150.

537. Rogoff, K. (1996) The Purchasing Power Parity Puzzle, Journal of Economic Literature, XXXI ~ (June), 647 - 668.

538. Roll, R. (1977) A Critique of Asset Pricing Theory's Tests, Journal of Financial Economics, 4, 1073 - 1103.

539. Roll, R. (1986) The Hubris Hypothesis of Corporate Takeovers, Journal of Business, 59(2), 541 - 566.

540. Roll, R. and Ross, S. A. (1980) An Empirical Investigation of the APT, Journal of Finance, 35(5), 1073 - 1103.

541. Roll, R. and Ross, S. A. (1984) A Critical Re - examination of the Empirical Evidence on the APT: A Reply, Journal of Finance, 39(2), 347 - 350.

542. Rouwenhorst, K. G. (1998) International Momentum Strategies, Journal of Finance, 53, 267 - 284.

543. Roy, A. D. (1952) Safety First and the Holding of Assets, Econometrica, 20, 431 - 449.

544. Rudebusch, G. D. (1995) Federal - Reserve Interest - rate Targeting, Rational Expectations, and the Term Structure, Journal of Monetary Economics, 35(2), 245 - 274.

545. Samuelson, P. (1963) Risk and Uncertainty: A Falacy of Large Numbers, Scienta, 98, 108 - 113.

546. Samuelson, P. (1965) Proof that Properly Anticipated Prices Fluctuate Randomly, Industrial Management Review, 6, 41 - 49.

547. Samuelson, P. A. (1969) Lifetime Portfolio Choice Selection by Dynamic Stochastic Programming, Review of Economics and Statistics, 51(3), 239 - 246.

548. Sarno, L. and Taylor, M. P. (1998) Real Exchange Rates under the Recent Float: Unequivocal Evidence of Mean Reversion, Economics Letters, 60, 131 - 137.

549. Sarno, L. and Taylor, M. P. (2002) The Economics of Exchange Rates, Cambridge University Press, Cambridge.

550. Sarno, L. and Thornton, D. L. (2003) The Dynamic Relationship between the Federal Funds Rate and the Treasury Bill Rate: An Empirical Investigation, Journal of Banking and Finance, 27, 1079 - 1110.

551. Sarno, L. and Valente, G. (2004a) Empirical Exchange Rate Models and Currency Risk: Some Evidence from Density Forecasts, Discussion Paper, University of Warwick.

552. Sarno, L. and Valente, G. (2004b) Modeling and Forecasting Stock Returns: Exploiting the Futures Market, Regime Shifts and International Spillovers, Journal of Applied Econometrics.

553. Savage, L. (1964) The Foundations of Statistics, John Wiley, New York.

554. Schwert, G. W. (1989) Why Does Stock Market Volatility Change Over Time? Journal of Finance, 44 (5), 1115 - 1153.

555. Scott, L. O. (1990) Asset Prices, Market Fundamentals, and Long - term Expectations: Some New Tests

of Present Value Models, Unpublished, University of Georgia, Athens, Ga.

556. Segal, U. (1989) Anticipated Utility: A Measure Representation Approach, Annals of Operations Research, 19, 359 - 373.

557. Sentana, E. and Wadhwani, S. (1992) Feedback Traders and Stock Return Autocorrelations: Evidence From a Century of Daily Data, Economic Journal, 102(411), 415 - 425.

558. Seo, B. (2003) Nonlinear Mean Reversion in the Term Structure of Interest Rates, Journal of Economic Dynamics and Control, 27(11/12), 2243 - 2265.

559. Sercu, P., Uppal, R. and vanHulle, C. (1995) The Exchange Rate in the Presence of Transaction Costs: Implications for Tests of Purchasing Power Parity, Journal of Finance, 50(4), pp. 1309 - 1319.

560. Sharpe, W. F. (1966) Mutual Fund Performance, Journal of Business, 39(1), 119 - 138.

561. Sharpe, W. F. (1975) Adjusting for Risk in Portfolio Performance Measurement, Journal of Portfolio Management, Winter, 29 - 34.

562. Shea, G. S. (1989) Ex - post Rational Price Approximations and the Empirical Reliability of the Present value Relation, Journal of Applied Econometrics, 4(2), 139 - 159.

563. Shea, G. S. (1992) Benchmarking the Expectations Hypothesis of the Term Structure: An Analysis of Cointegration Vectors, Journal of Business and Economic Statistics, 10(3), 347 - 365.

564. Sheffrin, S. M. (1983) Rational Expectations, Cambridge University Press, Cambridge, Mass.

565. Sherif, M. (1937) An Experimental Approach to the Study of Attitudes, Sociometry, 1, 90 - 98.

566. Shiller, R. J. (1979) The Volatility of Long - term Interest Rates and Expectations Models of the Term Structure, Journal of Political Economy, 87(6), 1190 - 1219.

567. Shiller, R. J. (1981) Do Stock Prices Move too Much to be Justified by Subsequent Changes in Dividends? American Economic Review, 71, 421 - 436.

568. Shiller, R. J. (1989) Market Volatility, MIT Press, Cambridge, Mass.

569. Shiller, R. J. (1990) Speculative Prices and Popular Models, Journal of Economic Perspectives, 4(2), 55 - 65.

570. Shiller, R. J. (1999) Human Behavior and the Efficiency of Financial Markets, Handbook of Macroeconomics, 1, 1305 - 1340.

571. Shiller, R. J. and Beltratti, A. E. (1992) Stock Prices and Bond Yields, Journal of Monetary Economics, 30, 25 - 46.

572. Shiller, R. J. and Perron, P. (1985) Testing the Random Walk Hypothesis: Power Versus Frequency of Observation, Economics Letters, 18, 381 - 386.

573. Shleifer, A. and Summers, L. H. (1990) The Noise Trader Approach to Finance, Journal of Economic Perspectives, 4(2), 19 - 33.

574. Shleifer, A. and Vishny, R. W. (1990) Equilibrium Short Horizons of Investors and Firms, American Economic Review, 80(2), 148 - 153.

575. Shleifer, A. and Vishny, R. W. (1997) The Limits of Arbitrage, Journal of Finance, 52(1), 35 - 55.

576. Siegel, J. (1998) Stocks for the Long Run, 2nd edition, McGraw - Hill, New York.

577. Simon, D. P. (1989) Expectations and Risk in the Treasury Bill Market: An Instrumental Variables Approach, Journal of Financial and Quantitative Analysis, 24(3), 357 - 365.

578. Simons, K. (1999) Should US Investors Invest Overseas? New England Economic Review (Federal Reserve Bank of Boston), Nov. /Dec., 29 - 39.

579. Singleton, K. J. (1980) Expectations Models of the Term Structure and Implied Variance Bounds, Journal of Political Economy, $ 8(6), 1159 - 1176.

580. Sirri, E. R. and Tufano, P. (1993) Buying and Selling Mutual Funds: Flows, Performance, Fees and Services, Working Paper, Harvard Business School, Cambridge, Mass.

581. Skinner, D. and Sloan, R. (2002) Earnings Surprises, Growth Expectations and Stock Returns, or, Don't Let an Earnings Torpedo Sink Your Portfolio, Review of Accounting Studies, 7, 289 - 312.

582. Smith, P. N. (1993) Modeling Risk Premia in International Asset Markets, European Economic Review, 37(1), 159 - 176.

583. Smith, P. N. and Wickens, M. R. (2002) Asset Pricing with Observable Stochastic Discount Factors, Journal of Economic Surveys, 16, 397 - 446.

584. Smith, P. N., Sorensen, S. and Wickens, M. R. (2003) Macroeconomic Sources of Equity Risk, Working Paper, University of York.

585. Solnik, B. (1974) Why not Diversify Internationally Rather than Domestically, Financial Analysts Journal, Jul. /Aug., 48 - 54.

586. Soros, G. (1987) The Alchemy of Finance, Simon and Schuster, New York.

587. Stamburgh, R. F. (1999) Predictive Regressions, Journal of Financial Economics, 54, 375 - 421.

588. Sullivan, R., Timmermann, A. and White, H. (1999) Data - snooping, Technical Trading Rule Performance, and the Bootstrap, Journal of Finance, 65(5), 1647

589. Sullivan, R., Timmermann, A. and White, H. (2001) Dangers of Data Mining: The Case of Calendar Effects in Stock Returns, Journal of Econometrics, 105(1), 249 - 286.

590. Summers, L. H. (1986) Does the Stock Market Rationally Reflect Fundamental Values? Journal of Finance, 41(3), 591 - 601.

591. Swensen, D. (2000) Pioneering Portfolio Management, The Free Press, New York.

592. Taylor, A. M. (2001) Potential Pitfalls for the Purchasing Power Parity Puzzle? Sampling and Specification Biases in Mean Reversion Tests of the Law of One Price, Econometrica, 69(2), 473 - 498.

593. Taylor, M. P. (1987) Covered Interest Parity: A High - frequency, High Quality Data Survey, Economica, 54, 429 - 438.

594. Taylor, M. P. (1988) What Do Investment Managers Know? An Empirical Study of Practitioners Predictions; Economica, 55, 185 - 202.

595. Taylor, M. P. (1989a) Covered Interest Arbitrage and Market Turbulence, Economic Journal, 99(396), 376 - 391.

596. Taylor, M. P. (1989b) Expectations, Risk and Uncertainty in the Foreign Exchange Market: Some Results Based on Survey Data, Manchester School, 57(2), 142 - 153.

597. Taylor, M. P. (1989c) Vector Autoregressive Tests of Uncovered Interest Rate Parity with Allowance for Conditional Heteroscedasticity, Scottish Journal of Political Economy, 36(3), 238 - 252.

598. Taylor, M. P. (1992) Modelling the Yield Curve, Economic Journal, 102(412), 524 - 537.

599. Taylor, M. P. and Peel, D. A. (2000) Nonlinear Adjustment, Long - run Equilibrium and Exchange Rate Fundamentals, Journal of International Money and Finance, 19, 33 - 53.

600. Taylor, M. P. and Sarno, L. (1998) The Behaviour of Real Exchange Rates During the Post - Bretton Woods Period, Journal ofInternational Economics, 46, 281 - 312.

601. Taylor, M. P., Peel, D. A. and Sarno, L. (2001) Nonlinear Mean - reversion in Real Exchange Rates: Towards a Solution to the Purchasing Power Parity Puzzles, International Economic Review, 42, 1015 - 1042.

602. Thaler, R. H. and Johnson, E. J. (1990) Gambling with the House Money and Trying to Breakeven - The Effects of Prior Outcomes on Risky Choice, Management Science, 36(6), 643 - 660.

603. Thaler, R. H., Tversky, A., Kahneman, D. and Schwartz, A. (1997) The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test, Quarterly Journal of Economics, 112, 647 - 661.

604. Timmermann, A. (1993) How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices, Quarterly Journal of Economics, 108, 1135 - 1145.

605. Timmermann, A. (1994) Why do Dividend Yields Forecast Stock Returns? Economic Letters, 46, 146

-158.

606. Timmermann, A. (1996) Excess Volatility and Predictability of Stock Returns in Autoregressive Dividend Models with Learning, Review of Economic Studies, 63(3), 523-558.

607. Timmermann, A. (2001) Structural Breaks, Incomplete Information and Stock Prices, Journal of Business and Economic Statistics, 19(3), 299-314.

608. Tirole, J. (1985) Asset Bubbles and Overlapping Generations, Econometrica, 53(5), 1071-1100.

609. Tobin, J. (1956) The Interest-elasticity of Transactions Demand for Cash, Review of Economics and Statistics, 38(3), 241-247.

610. Treynor, J. L. and Mazuy, K. (1996) Can Mutual Funds Outguess the Market? Harvard Business Review, 44, 131-136.

611. Treynor, J. L. (1965) How to Rate Management of Investment Funds, Harward Business Review, 43(1), 63-75.

612. Tzavalis, E. and Wickens, M. (1995) The Persistence of Volatility in the US Term Premium 1970-1986, Economic Letters, 49(4), 381-389.

613. Tzavalis, E. and Wickens, M. R. (1997) Explaining the Failure of the Term Spread Models of the Rational Expectations Hypothesis of the Term Structure, Journal of Money, Credit and Banking, 29(3), 364-380.

614. Vasicek, O. A. (1977) An Equilibrium Characterization of the Term Structure, Journal of Financial Economics, 5(2), 177-188.

615. Veronesi, P. (1999) Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model, Review of Financial Studies, 12, 975-1007.

616. Veronesi, P. (2003) The Peso Problem Hypothesis and Stock Market Returns, Journal of Economic Dynamics and Control, 28, 707-725.

617. Viceira, L. M. (2001) Optimal Portfolio Choice for Long-horizon Investors with Nontradable Labor Income, Journal of Finance, 56(2), 433-470.

618. Volkman, D. and Wohar, M. (1995) Determinants of Persistence in Relative Performance of Mutual Funds, Journal of Financial Research, XVIII, 415-430.

619. Walsh, C. E (1998) Monetary Theory and Policy, Princeton University Press.

620. Warther, V. A. (1995) Aggregate Mutual Fund Flows and Security Returns, Journal of Financial Economics, 39, 209-235.

621. Weil, P. (1992) Equilibrium Asset Prices with Undiversifiable Labour Income Risk, Journal of Economic Dynamics and Control, 16(3/4), 769.

622. Welch, I. (2000) Views of Financial Economists on the Equity Risk Premium and Other Issues, Journal of Business, 73, 501-537.

623. Welch, I. (2001) The Equity Premium Consensus Forecast Revisited, Working Paper, Yale School of Management.

624. Wermers, R. (2000) Mutual Fund Performance: An Empirical Decomposition into Stock-picking Talent, Style, Transactions Costs and Expenses, Journal of Finance, 55(2), 1655-1695.

625. West, K. D. (1987a) A Specification Test for Speculative Bubbles, Quarterly Journal of Economics, 102(3), 553-580.

626. West, K. D. (1987b) A Standard Monetary Model and the Variability of the Deutgchemark-dollar Exchange Rate, Journal of International Economics, 23, 57-76.

627. West, K. D. (1988a) Asymptotic Normality, When Regressors have a Unit Root, Econometrica, 56(6), 1397-1417.

628. West, K. D. (1988b) Dividend Innovations and Stock Price Volatility, Econometrica, 56, 37-61.

629. White, H. (1980) A Heteroscedasticity-consistent Covariance Matrix Estimator and a Direct Test for

Heteroscedasticity, Econometrica, 48, 55 – 68.

630. White, H. (1984) Asymptotic Theory for Econometricians, Academic Press, New York.

631. WM Company. (1999) Comparison of Active and Passive Management of Unit Trusts for Virgin Direct Financial Services.

632. Womack, K. L. (1996) Do Brokerage Analysts Recommendations Have Investment Value? Journal of Finance, 51(1), 137 – 167.

633. Woodford, M. (2003) Interest and Prices: Foundations of a Theory of Monetary Policy, Princeton University Press, Princeton, N. J.

634. Wurgler, J. and Zhuravskaya, K. (2002) Does Arbitrage Flatten Demand Curves for Stocks?, Journal of Business, 75, 583 – 608.

635. Xia, T. (2001) Learning about Predictability: The Effects of Parameter Uncertainty on Dynamic Asset Allocation, Journal of Finance, 56(1), 205 – 246.

636. Zheng, L. (1999) Is Money Smart? A Study of Mutual Fund Investors' Fund Selection Ability, Journal of Finance, 54(3), 901 – 933.

推荐读物

我们在这里精心选择了一个系列作为“推荐读物”。我们认为，金融经济学领域的这12本“最为满意的”书是对第二版《数量金融经济学》的有益补充，这些书对硕士和博士而言也是非常有用的。有些书关注的是金融经济学，有些书关注的是计量经济学，其他书会将你带入衍生产品和金融工程领域。所有书都值得你细细品味。

Campbell, J. Y., Lo, A. W. and MacKinlay, A. C. (1997) The Econometrics of Financial Markets, Princeton University Press.

这本书仍然是这一领域的经典参考文献。

Cochrane, J. H. (2001) Asset Pricing, Princeton University Press.

这本书是资产定价领域的经典理论基础，作者娴熟地将定价理论与相关计量经济技巧结合在一起。

Cerny, A. (2004) Mathematical Techniques in Finance, Princeton University Press.

这本书清晰地介绍了在实践中使用的衍生产品及套利的数学方法。

Cuthbertson, K. and Nitzsche, D. (2001) Investments Spot and Derivative Markets, John Wiley.

这本书是对金融市场的一个非常好的引论。

Cuthbertson, K. and Nitzsche, D. (2001) Financial Engineering: Derivatives & Risk Management, John Wiley.

这本书着重关注衍生产品的定价及其在风险管理中的应用。

网页(www.wiley.co.uk/cuthberston)上有以上两本书的幻灯片、Excell和Gauss文档及其他学生读物。

Hamilton, J. D. (1994) Time Series Analysis, Princeton University Press.

这是一本“圣经”，但要花费很多英镑。

Hull, J. C. (2003) Options, Futures and Other Derivatives; Prentice Hall.

这本书是衍生产品的简明而又易学的教材。

Luenberger, D. G. (1998) Investment Science, Oxford University Press.

这本书将现货和衍生资产理论的直观思想和例子漂亮地结合在一起。

Patterson, K. (2000) An Introduction to Applied Econometrics, Palgrave.

这本书是现代时间序列计量经济学的一个清晰导论。

Sarno, L. and Taylor M. P. T. (2002) The Economics of Exchange Rates, Cambridge University Press.

这本书覆盖的内容简明、全面、新颖。

Shiller, R. J. (2001) Irrational Exhuberance, Princeton University Press.

这本书是股票市场投资的经典叙述。

他写的所有东西阅读起来都非常简单(www. econ. yale. edu)。

Sydsaeter, K. and Hammond, P. (2002) Essential Mathematics for Economic Analysis, Prentice Hall.

这本书对基础数学引言进行了很好的叙述。